HOW DOES HIGH STAKES TESTING INFLUENCE TEACHERS' CLASSROOM INSTRUCTION?: INSTITUTIONAL PRESSURES AND CLASSROOM INSTRUCTION

by

Mika Yoder Yamashita

Yamanashi University, B.A., 1990

Lesley College, M.A., 2000

Submitted to the Graduate Faculty of

School of Education in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH

SCHOOL OF EDUCATION

This dissertation was presented

by

Mika Yoder Yamashita

It was defended on

February 3, 2011

and approved by

Maureen McClure, Ph.D., Associate Professor, Administrative and Policy Studies

Michael Gunzenhauser, Ph.D., Associate Professor, Administrative and Policy Studies

Cynthia E. Coburn, Ph.D., Associate Professor, Policy, Organization, Measurement and Evaluation, University of California, Berkeley

Dissertation Advisor: Mary Kay Stein, Ph.D., Professor, Learning Policy Center

HOW DOES HIGH STAKES TESTING INFLUENCE TEACHERS' CLASSROOM INSTRUCTION?: INSTITUTIONAL PRESSURES AND CLASSROOM INSTRUCTION

Mika Yoder Yamashita Ph.D.

University of Pittsburgh, 2011

This study examined how a total of eight math and science elementary school teachers changed their classroom instruction in response to high stakes and low stakes testing in one school district. The district introduced new assessment in the school year of 2005-06 to meet the requirement set forth by the No Child Left Behind Act (NCLB)—that the assessment should be aligned with the state academic standards. I conducted interviews with teachers and school administrators at two elementary schools, district officials, and a representative of a non-profit organization during the school year 2007-08 to examine how the new assessment introduced in 2005-06 had shaped classroom instruction.

Concepts from New Institutional Theory and cognitive approaches to policy implementation guided the design of this study. This study focused on how materials and activities associated with high stakes testing promoted ideas about good instruction, and how these ideas were carried to teachers. The study examined how teachers received messages about instruction and how they responded to the messages.

The study found that high stakes testing influenced teachers‘ classroom instruction more than low stakes testing; however, the instructional changes teachers made in response to state testing was at the content level. The teachers‘ instructional

strategies did not change. The teachers‘ instructional changes varied with the degree of implementation of existing math curriculum and with the degree of support they received in understanding the meaning of assessment results.

The study concluded that, among the six teachers I studied, high stakes testing was not a sufficient intervention for changing teachers‘ instructional strategies. The study also addressed the challenges of aligning instructional messages across assessment, standards, and curriculum.

TABLE OF CONTENTS

1.0 INTRODUCTION 1

2.0 REVIEW OF LITERATURE 5

2.1 INTRODUCTION 5

2.2 DOES HIGH STAKES TESTING CHANGE CLASSROOM INSTRUCTION MORE THAN LOW STAKES TESTING? 7

2.2.1 Studies that Used Teacher Self-Report 7

2.2.2 Studies that Used Observation 9

2.3 WHAT INSTRUCTIONAL CHANGES DO TEACHERS MAKE? 11

2.3.1 Studies that Used Teacher Self-Report 11

2.3.2 Studies that Used Observation 15

2.4 WHAT INQUIRIES ARE NEEDED? 19

2.4.1 How Does A School's Organizational Context Shape Instructional Change? 20

2.4.2 How Do Different Academic Subjects Shape Instructional Change? 20

2.4.3 How Do Policy Pressures from Testing Shape Instructional Change in the Long Term? 21

2.4.4 What Messages Do Teachers Come Across? 22

2.5 SUMMARY 23

3.0 CONCEPTUAL FRAMEWORK 24

3.1 INTRODUCTION 24

3.2 NEW INSTITUTIONAL THEORY 25

3.3 THEORETICAL DEVELOPMENT OF NEW INSTITUTIONAL THEORY 27

3.4 COGNITIVE APPROACH TO POLICY IMPLEMENTATION 32

3.5 CONCEPTUALIZATION OF THIS STUDY 36

4.0 RESEARCH QUESTIONS 42

5.0 METHODS 44

5.1 DESIGN OF THE STUDY 44

5.2 DATA COLLECTION AND ANALYSIS FOR RESEARCH QUESTION 1: CHANGE OF INSTITUTIONAL LOGICS 46

5.3 DATA COLLECTION AND ANALYSIS FOR RESEARCH QUESTION 2: CHANGE OF INSTRUCTION 47

5.3.1 Data Collection 47

5.3.2 Data Analysis 54

5.4 GENERALIZABILITY, VALIDITY AND LIMITATIONS OF THE STUDY 58

5.4.1 Generalizability 58

5.4.2 Validity 59

5.4.3 Limitations of the Study 61

6.0 FINDINGS: CHANGE OF ORGANIZATIONAL FIELD 64

6.1 INTRODUCTION 64

6.2 FROM 1989 TO 1995: PROMOTING REFORM-ORIENTED INSTRUCTION 68

6.2.1 Institutional Logic: Reform-Oriented Instruction 68

6.2.2 Governance 73

6.2.3 Spread of Institutional Logics 75

6.3 FROM 1996 TO 2003: PROMOTION OF REFORM ORIENTED INSTRUCTION AND TEACHING FOR THE TEST 76

6.3.1 Institutional Logics: Reform-Oriented Instruction and Teaching for the Test 77

6.3.2 Governance 86

6.3.3 Spread of Institutional Logics 87

6.4 FROM 2004-2008: MOVING FROM REFORM-ORIENTED INSTRUCTION TO TEACHING FOR THE TEST 90

6.4.1 Institutional Logics: From Reform Oriented Instruction to Teaching for the Test.....91

6.4.2 Governance 108

6.4.3 Spread of Institutional Logics 110

6.5 SUMMARY 112

7.0 FINDINGS: POLICY PRESSURES ASSOCIATED WITH MATH AND SCIENCE IN THE FIELD OF DCPS 114

7.1 INTRODUCTION 114

7.2 DIFFERENCE IN POLICY PRESSURES 115

7.3 DIFFERENCE IN MATH AND SCIENCE INSTRUCTIONAL POLICY 117

7.4 SUMMARY 118

8.0 FINDINGS: SCHOOL CONTEXT 120

8.1 INTRODUCTION: TWO ELEMENTARY SCHOOLS 120

8.2 JEFFERSON ELEMENTARY SCHOOL 120

8.3 CENTRAL ELEMENTARY SCHOOL 127

8.4 SUMMARY 131

9.0 FINDINGS: CHANGE OF MATH INSTRUCTION 133

9.1 INTRODUCTION 133

9.2 THREE TEACHERS' CHANGE OF INSTRUCTION 134

9.2.1 Mrs. Sommers at Central Elementary School 135

9.2.2 Mrs. Green at Central Elementary School 146

9.2.3 Mrs. Canright atJefferson Elementary School 156

9.3 ANALYSIS 167

9.3.1 Nature of Policy Pressure 169

9.3.2 Organizational Context 187

9.3.3Teachers‘ Knowledge 202

9.4 SUMMARY 207

10.0 FINDINGS: CHANGE OF SCIENCE INSTRUCTION 209

10.1 THREE TEACHERS' CHANGE OF INSTRUCTION 210

10.1.1Mr. Shawver at Jefferson Elementary School 210

10.1.2 Mr. Berry at Central Elementary School 217

10.1.3Mrs. Green at Central Elementary School 226

10.2ANALYSIS 231

10.2.1 Policy Signal 232

10.2.2 Organizational Context 241

10.2.3 Teacher‘s Knowledge 243

10.2.4 Summary 245

11.0 ROLE OF HIGH STAKES TESTING FOR CHANGING INSTRUCTION WITHIN THE TEST BASED ACCOUNTAIBLITY POLICY CONTEXT 246

11.1 LOW AND HIGH STAKES TESTING AND INSTRUCTIONAL CHANGE 247

11.2 INSTRUCTIONAL CHANGE BY HIGH STAKES TESTING 250

11.3 STANDARDS, CURRICULUM, AND ASSESSMENT'S INFLUENCE IN CHANGING TEACHERS' CLASSROOM INSTRUCTION 255

11.4 CHANGE OF INSTRUCTIONAL PRACTICE IN STANDARDS-BASED AND TEST-BASED ACCOUNTABILITY CONTEXT 261

11.5 SUMMARY 263

12.0 SUMMARY AND CONCLUSION 265

12.1 THIS STUDY AND NEW INSTITUTIONAL THEORY 265

12.2 THIS STUDY AND STUDIES OF THE IMPACT OF HIGH STAKES TESTING ON CLASSROOM INSTRUCTION 269

12.3 AREAS OF FURTHER STUDIES 271

APPENDIX A INTERVIEW PROTOCOL FOR DISTRICT OFFICIALS274

APPENDIX B INTERVIEW PROTOCOL FOR TEACHERS 277

APPENDIX C CLASSROOM OBSERVATION PROTOCOL 283

APPENDIX D CODING STRUCTURE 285

BIBLIOGRAPHY 286

LIST OF TABLES

Table 1 School Demographics 51

Table 2 Historical AYP Report 51

Table 3 Data Collection from Teachers 53

Table 4 Change of DC Public Schools' Organizational Field 67

Table 5 Timetable of Policy Initiatives between 1989 and 1996 69

Table 6 Timetable of Policy Initiatives between 1996 and 2004 77

Table 7 Timetable of Policy Initiatives between 2004 and 2008 91

Table 8 Policy Pressures Associated with DC CAS 115

Table 9 List of Policy Pressures and Teachers' Responses 172

Table 10 List of Policy Pressures and Teachers' Response 234

LIST OF FIGURES

Figure 1 Conceptual Framework of This Study..... 37

1.0 INTRODUCTION

The federal No Child Left Behind (NCLB) Act of 2001 required states to administer high stakes testing to illustrate improvement in students‘ academic performance. Building upon the approach taken by previous federal laws to improve teaching and learning, NCLB expanded the accountability function of state assessment. Federal laws since the 1990s, including NCLB, are based on the assumption that teaching and learning will improve by aligning state standards, curriculum, professional development, and assessment. These federal policies envisioned two functions of state assessment for improving teaching and learning. First, the state test was meant to reflect standards with the goal of helping teachers change their instruction in line with those standards. Second, state assessment intended to illuminate a school‘s level of accountability for improving student academic performance (Carnoy, Elmore & Siskin, 2003; West & Peterson, 2003).

NCLB expanded the second function of state assessment by publishing each individual school‘s performance on the state test for each tested grade and by prescribing specific consequences for the schools that did not meet with the adequate yearly progress standard on the state test. Attaching high stakes to the assessment was ―an incentive for low performing schools to improve" because ―schools that want to avoid losing students, not to mention restructuring, will have to do a better job" (Department of Education,

2002, p. 11). As it is, the law is based on the assumption that high stakes will make teachers teach for standards, and improve teaching and learning at all schools.

My study examines the extent to which this policy assumption is true. In other words, it examines how teachers changed their classroom instruction in response to high stakes testing. Although there are studies about the impact of external testing on classroom instruction, there are only a small number of studies that examine the role of high stakes testing aligned with state standards and curriculum.

I conducted my study in one school district, the District of Columbia Public Schools (DCPS). DCPS implemented new standards and district assessment to comply with NCLB in 2005. DCPS had administered high stakes testing since the late 1990s, but the test was not aligned with the standards. NCLB required the district to create assessment that aligned with the district‘s standards. In order to understand how high stakes testing changed teachers‘ classroom instruction, I collected data from a total of eight math and science teachers at two elementary schools during the school year of 2007-08. I asked teachers how their instruction had changed over time in order to understand the ways high stakes testing shaped their instruction, if at all. Rather than assuming there was a direct relationship between high stakes testing and changes in instruction, I collected data to understand more about *how* they had changed instruction within a complex school context. I collected data to discover how their instructional change in response to high stakes testing was different from instructional change in response to other policy pressures. In short, I collected data from the teachers‘ perspectives.

I found that the teachers thought high stakes testing had influenced their instruction, and that the teachers changed instruction so that they could cover the content that would be tested. Although the high stakes testing influenced the content of instruction, it rarely changed pedagogy. A majority of the time, I observed the teachers teaching for the test in a traditional instructional approach that was not in line with the instructional approach promoted by the district's standards and curriculum. Based on the data collected from six elementary school teachers, my study contributes to prior field investigations, concluding that high stakes testing does not change teachers' pedagogy in line with the standards.

This study draws conceptual tools from New Institutional Theory and Cognitive Approaches to Implementation Study. These approaches allowed me to examine how policy messages originated at the district were carried to teachers and how teachers interpreted these messages. They allowed me to examine what constituted high stakes testing policy, and how teachers understood this policy and made decisions about instruction. They also allowed me to consider how school contexts shaped the way teachers changed instruction. By taking into account the interpretations of policy messages and the role of context, this study was able to document how teachers enacted high stakes testing policy in their classrooms in different ways.

This type of inquiry is important because it examines the process through which teachers change classroom instruction. The approach I took in this study allowed me to describe the instructional change within the complex policy and school contexts in which everyday instruction occurred. I found that although all teachers were aware of and concerned about their students' academic performance on the high stakes tests, they

incorporated suggestions and guidance associated with the test differently. Hence, I concluded that high-stakes testing was not a sufficient lever to change these teachers' classroom instruction in line with state standards.

Below I present my study in the following order. In Chapter Two, I offer a review of existing literature on the influence of external testing on classroom instruction to discuss how the design of my study expands our current understandings. In Chapter Three, I outline and explain some conceptual tools I use from New Institutional Theory and Cognitive Approaches to examine how these tools broaden and complicate the analysis of the role of high stakes testing on classroom instruction. In Chapter Four, I present the following research questions that guide this study.

1. How did the institutional logics change as DCPS initiated new instructional and assessment policies in response to NCLB?
2. How did teachers change their classroom instruction?

In Chapter Five, I describe my methods of data collection and analysis. In Chapters Six through Eleven, I present my findings. To understand the role of high stakes testing in shaping teachers' classroom instruction, I describe how teachers changed their classroom instruction in response to math high stakes testing and science low stakes testing. In Chapter Twelve, I illustrate how the empirical evidence I collected in this study enriches theoretical discussions in New Institutional Theory and expands existing knowledge about the impact of high stakes testing on classroom instruction.

2.0 REVIEW OF LITERATURE

2.1 INTRODUCTION

A growing body of studies examine the impact of external testing on classroom instruction. Findings from these studies present an inconclusive picture about the role of high stakes testing in teachers‘ instructional change. Based on the existing studies I discuss, I find there is a need for studies that examine the process through which policy pressures associated with high stakes testing influence teachers‘ classroom instruction.

As my study focuses on the role of high stakes testing on changing teachers‘ classroom instruction, I review existing studies that are centered around the following two questions: a) Does high stakes testing change classroom instruction more than low stakes testing? b) What are the instructional changes teachers make in response to high stakes testing?

I explain how existing studies answer each of these questions. I present these studies according to the methods they use because the methods directly affect a particular study‘s conclusions (Cimbricz, 2002). First, I review studies that use teachers‘ self-

reports as data sources. Second, I review studies that incorporate observation. After describing what existing studies reveal about these two questions, I discuss what inquiries are needed to layer our understandings about the role of high stakes testing in changing teachers' classroom instruction.

In this review, I define high stakes testing as a testing program whose results have consequences for teachers or students (i.e., test results determine whether students can graduate, or test results determine whether schools should be reconstructed), and I define low stakes testing as testing that does not have such consequences for teachers or students. For low stakes testing, results are provided to teachers or to the public, but no direct consequences are attached to the results[1].

The studies I review include different types of assessment, although not all studies provide a complete description about the stakes attached and nature of the assessment. Some studies examine state assessments that are comprised of multiple-choice questions. Often this type of assessment intends to measure mastery of certain knowledge or skills. Other state assessments include portfolio assessment or group activities. The focus of

[1] Researchers differently define what entails high and low stakes. Madaus (1988) defines the high and low stakes testing from the perceptions of teachers, principals, or students. She defines high stakes as meaning "whose (tests') results... are seen by students, teachers, administrators, parents, or the general public, as being used to make important decisions that immediately and directly affect them" (Madaus, 2003, p. 87) . Low stakes "is perceived as not having important rewards or sanctions tied directly to test performance" (Madaus, 2003, p. 88). Heubert and Hauser (1999) distinguish "high" from "low" by focusing on assumptions about how to change people's action held by policy makers. They define "low stakes testing (as having) no significant, tangible, or direct consequence attached to the results with information alone assumed to be sufficient incentive for people to act. In contrast, high stakes policies assume that information alone is insufficient to motivate educators to teach well and students to perform to high standards. Hence, it is assumed the promise of rewards or the threat of sanctions is needed to ensure change" (pp. 36-7). Abrams, Pedulla Madaus (2003) categorizes stakes into high, moderate and low, and they categorize publishing test results as moderate stakes. In this review I followed Heubert and Hauser's definition.

these assessments seeks to measure conceptual understanding and problem solving skills. Some assessments are national norm tests and not in line with state standards, while other assessments are in line with state standards. I include descriptions about the nature of assessment as much as possible to show what we already know about the relationship between the nature of high stakes testing policy and instructional change.

2.2 DOES HIGH STAKES TESTING CHANGE CLASSROOM INSTRUCTION MORE THAN LOW STAKES TESTING?

2.2.1 Studies that Used Teacher Self-Report

In general, studies that used teachers‘ self-reports through survey or interviews conclude from cross-state comparative studies that the higher the stakes attached to the test, the more instructional change that happens (Abrams, Pedulla &Maddans, 2003; Carnoy, Elmore & Siskin, 2003; Wilson & Corbett, 1991). A study of teachers in a low stakes testing state concludes that testing rarely influences instruction (Glasnappo, Paggio & Miller, 1991). A study of teachers at different grade levels found that teachers in testing grades are more likely to change their instruction (Stecher & Barron, 1999). Taken together, these studies present the image that the higher the stakes attached to the test, the more instructional change occurs.

However, a small number of studies raise questions about the above general conclusion. First, a careful reading of the studies suggests that because the studies are based on teachers' self-reports, reported instructional change could be conflated with teachers' perceptions of intrusiveness of these tests in their instruction. For example, in a study that compares the impact of high stakes testing in Maryland and low stakes testing in Maine, Wilson and Corbert (1991) found that the teachers in Maryland perceived recently implemented tests—citizenship and writing—as more intrusive to their instruction than the test of reading and mathematics that had been in place for a while. Hence, the instructional change reported by Maryland's teachers may be exaggerated, which suggests the need for triangulating teachers' reports with information from other sources.

Second, a small number of studies report that low stakes testing makes teachers teach for the test, although many of these studies do not compare the instructional changes with the ones for high stakes testing. (Firestone, Monfils & Schorr, 2004; Smith, 1991; Stecher & Mitchell, 1995). The teachers in these studies report that their instruction changed to be in line with state assessments. McDonnell and Choisser (1997) interviewed teachers in both a high stakes test state, Kentucky, and a low stakes test state, North Carolina[2], and they found that teachers in both states were taking the assessment equally seriously and aligning their instruction toward the test by such methods as teaching problems that have appeared on the test and changing the format of class

[2] When this study was conducted in 1993, teachers did not see the test as having any consequence for them because of the lack of funding and union resistance was strong. Studies conducted in the late 90s and after 2000 categorize tests in North Carolina as high stakes because the state established a system to sanction teachers and schools based on their schools' performance on the test in the 1996-1997 school year.

assignments to more directly reflect the state assessment. Smith (1991) reports that the publication of test results makes teachers feel guilty and ashamed; hence, teachers teach *to* the test. In summary, the studies that use teachers' self-reports are inconclusive as to whether or not high stakes testing actually changes instruction more than low stakes testing.

2.2.2 Studies that Used Observation

My literature search did not find studies that use observation of classroom instruction prior and after the implementation of testing programs. Among the studies I reviewed, there is only one cross-state study that presents how instruction differs between teachers in high stakes testing states and low stakes testing states[3].

Firestone and Mayrowetz (2000) compared classroom math instruction in states with low and high stakes testing, concluding that teachers in Maryland (high stakes state) are more likely to incorporate activities that appear on the state test than the teachers in Maine (low stakes state). Both states' math tests are in line with the state standards that aim at instruction promoted by the National Council of Teachers of Mathematics. The problems in both states' assessments are intended to promote problem solving activities and the conceptual understanding of mathematics.

---

[3] There is one more study that compares teachers in Maryland and Maine by Firestone, Mayrowetz and Fairman (1998). The finding is the same with the findings from Firestone & Mayrowetz (2000). Teachers in high stakes state made more changes in order, but instructional methods did not differ in both states.

Although this finding (that the higher the stakes, the more the instructional change) is in line with many studies that use teachers‘ self-report, Firestone and Mayrowetz (2000) conclude from their observation of lessons that ―external pressure is less effective in changing instructional strategy than content taught” (Firestone & Mayrowetz, 2000, p. 736). For example, they observe that teachers in Maryland assigned math activities from the state test, but ―the MSPASP activities we observed often proved flawed” (Firestone & Mayrowetz, 2000, p. 736). Although the intent of the assessment activity is to get students to explore and discover strategies, they observe that ―rather than encouraging students to explore various mathematical concepts to work on the program, teachers provided step-by-step procedures to solve problems?” (Firestone & Mayrowetz, 2000, p. 736).

In addition, Firestone and Mayrowetz (2000) found that teachers‘ instructional strategies, such as the size of the problem given to the students and patterns of interaction between teachers and students, are not notably different. The study suggests that high stakes testing is more likely to make teachers introduce activities or skills that are tested on the test, but the levels of stakes do not make much difference in changing teachers‘ instructional strategies.

2.3 WHAT INSTRUCTIONAL CHANGES DO TEACHERS MAKE?

2.3.1 Studies that Used Teacher Self-Report

The studies that use teachers‘ self-reports presents various instructional changes teachers made. In general, studies agree that the teachers change the content of instruction. Teachers cover the content that appears on the test and incorporate the tasks and activities that appear on the test. In addition, the studies find that the format of the test may shape the instructional change. Studies that use detailed questionnaires find that teachers‘ beliefs about student learning and pedagogy are less likely to be transformed. As it is, the studies suggest that instructional change in response to an external test is often at the content and activity level, and teachers‘ beliefs rarely change. Furthermore, one study that uses teachers‘ self reports shows different instructional changes among teachers in response to a state assessment, which suggests a need for further investigation of why individual teachers variably change instruction in response to the same testing program.

Studies vary in the instructional change examined. A majority of studies view instructional change in terms of change in allocation of time or change in content of instruction (Abrams, Pedulla & Maddans, 2003; Hargrove, et al, 2000; Moon, Brighton & Callahan, 2002; Wilson & Corbert, 1991). Teachers often respond that they teach tested content and subjects. Hence, many studies conclude that testing narrowed instruction, which projects a negative effect of testing on instruction.

There are some studies that project a positive effect. Teachers in Kentucky report that they became more focused on student achievement than before, and teachers in four

states, Kentucky, New York, Georgia, and Pennsylvania, report that their instruction became more focused than before (Hamilton, et al., 2007; Koretz, Barrons, McDonnell & Choisser, 1997; Mitchell & Stecher, 1996;). Teachers in Kentucky, which has high stakes portfolio assessment, report that they spend more time on writing and demonstrating how to perform new mathematical skills. They also report that they are giving more assignments that include writing (Stecher & Barron, 1999). These changes are in line with the state standards.

Researchers use various methods to capture instructional change. A number of studies look at the specific instructional activities teachers do in the classroom, often providing the instructional activities as selections teachers can check off in a survey questionnaire, such as using assessment that takes the same format as the state assessment, lecturing, facilitating discussion, or introducing activities that appear on the state test (McDonnell & Choisser, 1997; Stecher & Barron, 1999; Stecher & Mitchell, 1995). The researchers found that teachers introduce activities suggested in the state assessments. For example, from interviews with teachers in Kentucky and North Carolina, researchers find that teachers in North Carolina are more likely to use multiple-choice for their class assignments than teachers in Kentucky (McDonnell & Choisser, 1997). Taken together, these studies suggest that teachers tend to align their instruction to the tasks, skills, concepts, and format of assessment.

Several studies examine the extent to which high stakes testing gears classroom instruction toward the instruction promoted by state standards. They compare teachers‘ reported instructional change and the instructional activities envisioned in the state standards. In these studies, the researchers employ various questioning techniques to

determine if a teacher‘s classroom instruction is in line with the state standards. These techniques include: using a time log to accurately document how teachers spend time on specific instructional activities (McDonnell & Choisser, 1997) and asking what instructional approaches the teachers will take based on a hypothetical students‘ responses to a problem. These approaches inform teachers‘ subject knowledge and views of student learning (Hamilton,2004; Hamilton, Brends & Stecher, 2005; Koretz, Barron, Mitchell & Stecher,1996, Stecher & Mitchell, 1995). Thus, besides asking about the existence of specific instructional activities, these studies examine teachers‘ instructional and subject knowledge, and their beliefs about instruction.

These studies suggest that although teachers report that they change instruction, their belief about student learning and knowledge of instruction rarely changes. For example, although the teachers in Kentucky report that their views of subject knowledge and instruction are changing to the ones promoted by the state standards, which are based on the assumption that all students can learn to high standards, the researchers find that less than half of teachers expect that all students can learn, and 63-70 % of teachers agree to the statement that ―there are some students who will never perform above the level of novice, no matter how much effort the teacher devotes to them” (Stecher, Barron, Kaganoff & Goodwin, 1998, p. 24). This finding indicates that despite the promotion of the idea that all students can perform to high standards, teachers‘ predominant views do not change.

Studies that use cross-state analysis provide insight into factors that seem to affect how teachers change instruction. The studies find that the nature of assessment, the maturity of implementation, and the nature of policy instrument shape instructional

change. A cross-state study suggests that forms of testing may guide instructional change, i.e. portfolio assessments change classroom instruction to student centered and inquiry based, but multiple-choice testing narrows instruction (Abrams, Pedulla & Mandaus, 2003). Studies that examine portfolio assessment (KIRIS) in Kentucky and assessment that aims at progressive instruction in New Jersey are in line with this conclusion (Firestone, Montils, & Schorr, 2004; McDonnel & Choisser, 1997;).

Existence of support for implementing new state assessments, such as professional development, seems to influence how teachers change classroom instruction. Hamilton and colleagues (2005) find from a study of four states, Georgia, Pennsylvania, New York and California, that the teachers in the three states other than Pennsylvania were more likely to align content of instruction with the state standards. The researchers argue that this difference may derive from the existence of professional development in the three states, as well as the longer time the state tests have been in place in those states.

Although these studies suggest that the nature of assessment and existence of support for teachers to learn about state testing may influence instructional change, a study of North Carolina‘s high stakes testing conducted by Jones and colleagues suggests the need for further investigation about factors that shape individual teachers‘ instructional change (Jones, et. al. , 1999). In this case, the state assessment was in line with the state standards, which include mostly multiple-choice questions and a couple of open-ended questions (McDonnel & Choisser, 1997). Jones and colleagues find that teachers change instruction differently even when they are under the same state assessment policy. A state wide survey found that 13.9% of teachers indicate that they are using more inquiry projects, but 13% of teachers answer that they are using fewer.

10.8% of teachers indicate that they were using more lecturing, but 13% indicate they were using less. 32.5% said they were using more hands-on activities, 14.7% said they were using fewer (Jones, et. al.,1999). The study suggests that the format of testing may not be the only factor that shapes the way teachers change their classroom instructional change.

Carnoy and colleagues (2003) suggest that school contexts shape schools' responses to high stakes testing pressure. They find schools that have strong internal accountability and a shared organizational norm can endure the external pressure of improving academic performance more easily than the schools that lack capacity and a shared organizational norm (Carnoy, Elmore & Siskin, 2003).

The review of studies suggests a need to examine intervening factors that shape different instructional changes, particularly at the individual and school level. Studies that use observation focus on how individual teachers change classroom instruction and how school context shapes this process.

2.3.2 Studies that Used Observation

None of the studies I reviewed compare classroom instruction between pre- and post implementation of high stakes testing. The studies I reviewed examine teachers in one state (Anagnostopolous, 2005; Booher-Jennings, 2005; Firestone, Mayrowetz & Fairman 1998; Sloan, 2000; Watanabe, 2007; Wong, Anagnostopoulous, Ruthledge & Edwards, 2003;), and most studies compare observed instruction with state standards and/or

assessment to understand how state standards and assessment shape teachers‘ classroom instruction. In these studies, the researchers use observation to better understand the nature of instruction by documenting interactions between teacher and students and by documenting types of tasks teachers assign students to identify how classroom instruction differs from intended instruction.

The studies of multiple schools and teachers identify factors that shape instructional change among teachers, including a school‘s organizational context and teachers‘ beliefs about good instruction (Lipman, 2004). The studies that use observation found that the standards and assessment do not necessarily convey the same instructional messages, and the assessment has a limited impact on changing teachers‘ instruction to reform-oriented instruction that promotes student centered learning.

The researchers find that the range of content and depth of cognitive exploration students experience in the classroom narrows as teachers teach under high stakes testing pressure (Diamond, 2007; Watanabe, 2007;). At the same time, studies that use observation conclude that the impact of state testing for changing teachers‘ classroom instruction is limited (Diamond, 2007; Firestone, Montils & Schorr, 2004). Firestone and colleagues find, in their the study of the impact of low stakes math testing on middle school math instruction in New Jersey, that even when assessment is designed to promote conceptual understanding and problem solving in math, teachers tend to teach specific skills and procedures. They conclude that tests with open-ended questions encourage more challenging forms of teaching to the test, but that testing alone is not enough to bring about such changes” (Firestone, Montils & Schorr, 2004). Diamond‘s (2007) study of Chicago‘s high stakes test that was not in line with the state standards found that

although teachers teach the content of the tests, pedagogical change rarely happened in response to high stakes testing. Diamond (2007) reports that teachers did not see high stakes testing as the sole influence on instruction. Rather, teachers reported that their instruction is shaped by various sources, such as colleagues and principals, and high stakes test is merely one influence among many.

Studies that examine state standards and assessment find that teachers receive different instructional messages from state standards and state assessment in terms of the depth of exploration students should enact. When teachers encounter the gap between suggested instruction by the standards and assessment, they tend to align their instruction towards assessment rather than state standards. Watanabe (2007) found North Carolina's state reading assessment does not cover all parts of the state standards. While the standard "press[es] teachers to adopt a more progressive pedagogy," the test asks questions that measure the mastery of skills and content knowledge (Watanabe, 2007, p. 354). She observes that teachers emphasize literal interpretations of the text rather than facilitating class discussion to help students examine the connection between reading and their life experiences. Watanabe (2007) concludes, "many teachers feel it is necessary to prepare their students for the state standardized tests" (Watanabe, 2007, p. 384), rather than teaching in a way standards encourage. Wong and colleagues find that while state standards promote instruction in line with standards-based instruction, the district learning standards translate it into traditional instruction (Anagnostopoulos, 2005; Wong, Anagnostopoulos, Ruthledge & Edwards, 2003). In their studies, teachers' instruction was predominantly teacher-centered, which was in line with district guidance.

Studies that compare teachers at different schools or multiple teachers suggest

that a school‘s organizational context and teacher beliefs shape the way teachers change their instruction. From a cross-school study in Chicago elementary schools, Lipman (2004) finds that teacher‘s instructional change for test preparation is minimal at schools with a strong norm and mission of ―developing sophisticated literacy and love of reading” and schools that ―respect teachers‘ complex work” (Lipman, 2004, p. 149). In this school, the teachers use test preparation materials for only two weeks with the clear understanding that educational activity around these materials are quite different from what they want students to do (Lipman, 2004). On the contrary, the teachers at schools that do not have this organizational norm use the test preparation material to guide instruction throughout the school year.

Studies that examine multiple teachers present other factors that shape how individual teachers change classroom instruction. These studies conclude external testing is an ―uncertain lever” (Grant, 2001, p. 398) because teachers‘ instruction is shaped by various factors, such as teachers‘ views of the subject matter and learners (Grant, 2001; Zancanella, 1992), and teachers‘ past experience and status in schools (Zancanella, 1992). For example, Grant (2001) reports that one social studies teacher did not cut the content that is not tested on the Regent Test that is used for determining students‘ promotion. This teacher viewed it as important for students to learn this particular segment of history to understand the Civil Rights Movement, even if the particular material is not tested, and it makes his lesson schedule tight. Zancanella (1992) reports one teacher buffered the pressure from the state literature test. This teacher continued discussion-based reading instruction despite a state test that asks questions that measure students‘ ability to identify information from a text. Zancanella (1992) finds this

teacher‘s established status as a literature teacher within the school is what enabled her to push back against testing pressure. Other studies find that teachers change instruction in response to testing not because they believe teaching for the standards is good, but because of their hope of protecting students‘ self-esteem. In this case, teachers change instruction not because they believe it helps students learn better, but because they want students to feel successful (Ketter & Pool, 2001; Mathison & Freeman, 2003).

In summary, the studies suggest that a school‘s organizational norms, teachers‘ beliefs and knowledge, and the nature of tests shape teachers‘ instructional change. The studies also show that improving instruction is not the only reason to align instruction with assessment. The studies suggest teachers may sometimes receive conflicting instructional messages between standards and assessment, and assessment tended to make teachers teach in more traditional ways despite the state standards promotion of reform-oriented instruction.

2.4 WHAT INQUIRIES ARE NEEDED?

Existing studies suggest that external testing does influence classroom instruction, but the impact is limited to the change of content, order of instruction, and introduction of activities that are encouraged by the test. The existing studies are not clear-cut about whether high stakes testing changes teachers‘ instruction more than low stakes testing. The existing studies also identify some intervening factors that shape instructional

change; however, these studies seldom describe *how* they shape instruction. Based on the review, I argue the following inquiries are needed.

2.4.1 How Does A School's Organizational Context Shape Instructional Change?

Although the studies suggest that existing capacity and norms of schools shape how schools respond to high stakes testing pressure (Carnoy, Elmore & Siskin, 2003;Lipman, 2004), we still do not know how these organizational contexts shape individual teachers' instructional change. Existing studies suggest that teachers understand the same policy pressures in different ways because of their different beliefs and experiences (Hamilton, 2004; Cohen, 1990). Hence, one of the ways to investigate this issue is to examine how teachers view testing pressures and how organizational context shapes their views.

2.4.2 How Do Different Academic Subjects Shape Instructional Change?

The majority of existing studies examine teachers within the same academic subject (Grant, 2001; Ketter & Pool, 201; Mathison & Freeman, 2003); hence, we do not know much about how teachers of different academic subjects change their instruction in response to high stakes testing. Only one study tells us that teachers who teach math

make more change in instruction than writing teachers (Stecher & Barrons, 1999)[4]. Yet Siskin (2003) reports that teachers of different academic subjects face different types of challenges when high stakes testing policy is implemented because academic subjects differ in the degree of tightness of knowledge components, or, in other words, stability and solidity of sequence in curriculum. There is a need to examine how the different kinds of academic knowledge influence how teachers change instruction.

2.4.3 How Do Policy Pressures from Testing Shape Instructional Change in the Long Term?

Existing studies tend to focus on teachers' instructional change at the early stages of the implementation of high stakes testing. For example, Zancanella (2001) reports that a middle school literacy teacher who has established status in the school buffered the external testing pressure and taught as she believed during the first year, but we lack studies that examine the long-term impact of the testing policy. There is a need to examine how teachers' responses change in the long run. The following observation also suggests the need for a long-term study. Researchers find that teachers do activities that are promoted by the test, but the ways they introduce these activities in the classroom are not in line with the instruction promoted by the testing programs. The teachers argue that

[4] By change of instruction, this study defines "spending more time on skills that are on the test." Stecher & Barrons (1999) found from surveys and interviews in Kentucky that more testing grade math teachers reported that they teach skills that are on the test than non-testing grade teachers. But teachers who taught writing did not change their teaching.

the activities give students the opportunity to reason about math issues and develop their own ways to represent and solve problems (Firestone, Fitz & Broadfoot, 1999; Firestone & Mayrowetz, 2000). In another study of implementation of new instructional activities, researchers find that some activities provide opportunities for new interpretations of policy messages as teachers implement them, which may reshape teachers‘ existing knowledge thereby shaping instruction (Spillane, Reiser & Reimer, 2002). Taken together, these studies suggest that external testing pressures could become a lever for instructional change over time. There is a need to examine how teachers change instruction over time.

2.4.4 What Messages Do Teachers Come Across?

Some existing studies compare instructional change intended by standards and the actual instructional change that happened in the classroom. These studies identify intentions of testing policy at the state level by assuming the testing policy reflects state standards (Hamilton, Breadt & Stecher, 2005; Stecher & Barron, 1999; Stecher, Barron Kaganoff & Goodwin, 1988; Stecher & Mitchell, 1995, Hamilton, et al. 2007). A small number of studies directly examine the policy messages teachers come across. They suggest that teachers receive conflicting messages about instruction from the state standards, district guidelines, and state assessment (Anagnostopoulos, Ruthledge & Edwards, 2003; Watanabe, 2007; Wong, Anagnostopoulou, Ruthledge & Edwards, 2003).

Hence, in order to have a better understanding of the impact of high stakes testing policy on teachers‘ classroom instruction, there is a need to examine the change of policy messages across different levels of governance, and the nature of the policy messages teachers actually receive. The implementation studies find that the nature of policy pressures shapes the way teachers respond to policy messages that suggest changing classroom instruction (Coburn, 2001, 2004; Spillane, Reiser & Reimer, 2002). These studies suggest that rather than assuming teachers receive uniform policy messages, there is a need to examine how policy messages shape their responses.

2.5 SUMMARY

From the review of the existing studies, I argue that there is a need for studies that examine the process though which various policy pressures associated with high stakes testing influence teachers‘ instruction. There is a need to examine how individual teachers differently understand policy pressures and how organizational contexts influence their understanding of policy pressures and instructional decisions. There is a need for multilevel studies that examine how policy messages are translated and what policy pressures teachers understand, feel, or interpret. My study took up these inquiries by using concepts from New Institutional Theory and Cognitive Approach. In the next section, I present how concepts from these two approaches assist in examining the process through which high stakes testing influences teachers‘ classroom instruction.

3.0 CONCEPTUAL FRAMEWORK

3.1 INTRODUCTION

I draw on concepts from the following two complementary approaches: New Institutional Theory (Scott, 2001) and the Cognitive Approach to Implementation (Spillane, Reiser & Reimer, 2002) to frame this study. Concepts from these approaches provide ways to investigate why individual teachers respond differently to the same high stakes testing policy, which is one of the areas for further inquiry identified in the literature review. Both New Institutional Theory and Cognitive Approach investigate cognitive aspects of social systems and human behaviors. The concepts in New Institutional Theory enable me to examine underlying notions of educational policy, such as what a particular policy imagines is good instruction and how assessment should be used for improving student academic performance. They also provide analytical tools to better understand how underlying concepts are carried by specific policy pressures The concepts used in Cognitive Approach enable me to consider the change of individual teachers‘ instruction by focusing on their changing beliefs and assumptions about good instruction and how students learn. As it is, the concepts from these two approaches help me to study what instructional messages these policies convey, what teachers understand about them, and what beliefs and experiences shape the way teachers respond to policy messages. Rather

than assuming there is a direct relationship between the high stakes testing and the change of instruction, the concepts from these two approaches allow me to consider how policy portrays specific kinds of ideas about instruction and assessment, how teachers make instructional decisions in relation to their beliefs about instruction, and how different school contexts shape this process.

3.2 NEW INSTITUTIONAL THEORY

New Institutional Theory conceptualizes the change of a particular organization, such as a school, non-profit organization, or private company, in relation to the change of its organizational environment by using "organizational field" as the unit of analysis. The organizational field is a collection of organizations that "constitutes [a] recognizable area of institutional life" and includes "key suppliers, resource and product consumers, regulatory agencies, and other organizations that produce similar services or products" (DiMaggio & Powell, 1983, p. 184). This conceptualization is beneficial for examining policy pressures on schools because of the de-centralized nature of the educational field, in which various non-system actors play an important role in shaping instructional policies (Coburn, 2005; Cohen 1982). These non-system actors, such as intermediary organizations or non-profit organizations, often carry policy messages by providing educational materials and providing trainings. In my study, I define the organizational field of schools I study as including the district, intermediary organizations, organizations

involved in evaluating and proposing instructional policies, teachers' professional communities, parents, and any other social or familial communities that are involved in providing suggestions about education. Hence, my study views teachers as coming into contact with messages about instruction and assessment that originate from these various sources.

Institutional theorists conceptualized that activities, material resources, laws and relations in the organizational field are organized around "institutional logics." The institutional logics are the field's "cognitive maps of belief systems" (Scott, Ruff, Mendel & Carol, 2000, p. 20). Institutional logics "specify what goals or values are to be pursued within a field or domain and indicate what means for pursuing them are appropriate" (Scott, Ruff, Mendal & Carol, 2000, p. 171). Activities, regulations, and practices issued and exercised by actors in the organizational field also promote, enforce, carry out, and spread institutional logics in the field.

The concept of institutional logics is beneficial for my study because it allows me to investigate messages about good instruction and use of assessment produced by various policy pressures. In this study, I conceptualize the change of institutional logics as the change in understandings about how teachers should teach and how assessment should be used in relation to instruction, because assessment policy was proposed as one way to improve student academic performance. I argue that institutional logics in DCPS include the following aspects: a) strategies to improve student academic performance; b) good instruction, or how teachers should teach; and c) assessment's use and relationship to instruction. I include, as part of good instruction, the concept of roles teachers and students inhabit during teaching and learning and how students learn academic subjects. I

examine how policy pressures associated with assessment and instruction issued in response to federal NCLB promote ideas and practices about good instruction and the relationship between instruction and assessment.

Drawing from New Institutional Theory provides comprehensive descriptions about the change of policy because New Institutional Theory offers a way to consider the various policy pressures teachers are exposed to. Below I present how New Institutional Theory might be used to examine issues associated with educational policy and its influence on classroom instruction. Then, I discuss how I position my study within the theoretical discussion in New Institutional Theory.

3.3 THEORETICAL DEVELOPMENT OF NEW INSTITUTIONAL THEORY

New Institutional Theory emerged in the field of organizational studies. New Institutional Theory provided the following two aspects to organizational studies:

1) Focus on examining the relationship between an organization and its external environment; and
2) Focus on examining cognitive aspects of institutional pressures that shape organizational structure and behavior.

While many organizational researchers who employ an open-system approach focus on resources and regulations in a particular organization‘s environment for shaping organizational change, the new institutional theorists argue that other types of

institutional pressures, such as taken-for-granted ideas and ways of doing things shared in organizational environment, shape organizational change (Scott, 2001). According to new institutional theorists, organizations need to align their structure to reflect these social expectations in order to survive, regardless of how they improve their technical performance. Organizational researchers call this focus ―new" Institutional Theory in contrast to ―old" Institutional Theory. In contrast, Old Institutional Theory focused on the normative institutional pressures that emerge from within the organization, such as organizational norms, to understand how organization operates (Hirsch & Lounsbury, 1997). Scott's (1995) definition of institutions presents three elements of institutions, which are inclusive of both Old and New Institutional Theory.

> Institutions consist of cognitive, normative and regulative structures and activities that provide stability and meaning to social behavior. Institutions are transported by various carriers—cultures, structures, and routines—and they operate at multiple levels of jurisdiction. (p. 33)

New institutional theorists have used educational organizations as critical cases to show that cognitive elements of institutional pressure shape organizational change. They show that it is more important for schools and districts to align their organizational structures and practices with socially shared ideas about how the school or district should function, even when it could prevent the organizations from improving organizational efficiency (Meyer & Rowan, 1977). Rowan (1982) found that the increase of administrative positions in one school district was not based on the size of the district. But it was the response to socially shared ideas about how schools should be organized (Rowan, 1982). Another study found that while various levels of people in educational

systems (i.e., teachers, principals, and superintendents) agree on administrative policy, agreement about instruction rarely exists. Based on this finding, institutional theorists argue that although a school's administrative practice is in line with the institutional environment, classroom instruction is "decoupled from administrative structure" (Meyer & Rowan, 1977; Meyer, Scott & Deal, 1981).

Although the focus on cognitive institutional pressure opened up new areas of study to understand how schools operate, early new institutional theorists did not examine how external pressures influence practices within the school (i.e. classroom instruction). Early new institutional theorists conceptualized that institutional pressures do not penetrate into classroom practice. Rather, schools respond to the institutional pressures by adding new educational offices, programs, or libraries because responding to these types of institutional pressures affirms the social legitimacy of the school. Hence, classroom instruction is decoupled from institutional pressures (Meyer & Rowan, 1977). This discussion projects the image that instructional policy does not reach into the classroom.

This line of inquiry may have arisen from the policy context at that time. During the 1970s and the 1980s, the focus of educational policy was on input, i.e., providing resources for improving school facilities. Current educational policies focus on outcome (i.e., improving student academic performance) (Carnoy, Elmore & Siskin, 2003; Elmore, Abelman & Fuhrman, 1996; Mazzeo, 2001; West & Peterson, 2003). Hence, new institutional theorists argue that since the institutional environment has changed to one that requires schools to demonstrate improved student outcomes, teachers' classroom instruction is no longer buffered from policy pressure (Elmore, 1999-2000). However,

there is little empirical evidence to support this claim because we have a limited number of studies that directly examine the change of classroom instruction in relation to the change of policy context.

Recently, new institutional theorists began examining change of classroom instruction (Coburn, 2001, 2004). Coburn‘s studies present important findings that require re-conceptualizing the decoupling argument. They also suggest a need for further investigations to identify the mechanisms by which individuals respond to institutional pressures. Coburn (2004) found that ideas and practices promoted in a policy context reach teachers and that teachers incorporate some of these ideas and practices into their classroom instruction. Her study identifies three factors that shape individual teachers‘ responses to institutional pressures: a) the nature of policy messages, b) the degree of frequency of the exposure to the policy messages; and c) the teachers‘ pre-existing beliefs and practices (Coburn, 2004). Studies of implementation of instructional policies conducted during the 1990s also found that the degree of concreteness of policy messages influence teachers‘ response to policy pressures (Spillane & Jennings, 1997).

These findings are in line with the studies that examined organizations‘ responses to institutional pressures (Goodstein, 1994; Oliver, 1991). These studies theorize that organizations strategically respond to institutional pressures; thus, institutional pressures influence organizations‘ practices and structures. The studies suggest that the content of institutional pressures influence organizations‘ responses. Organizations tend to comply with institutional pressures when the message of the institutional pressure is in line with the organizations‘ values and practices. In addition, organizations tend to comply with institutional pressures if the organizations see complying with the pressure improves their

effectiveness. These arguments suggest that classroom instruction is not completely buffered from institutional pressures, and new instructional policy can reach teachers, but that teachers' responses vary. My study extends the following two areas of inquiry.

First, my study examines teachers' responses within a test-based accountability policy context. Adding to the argument by recent new institutional theorists that policy pressures reach teachers, my study presents *how* institutional pressures that are promoted by law (i.e., high stakes testing) influence individual teachers' behavior. Coburn's (2001, 2004) study focuses on the change of reading policy in California, which was promoted through new reading standards. The study was conducted prior to the No Child Left Behind Act, when teachers had not yet encountered strong test-based accountability policy. Her study found that many policy pressures are communicated as normative pressures that encourage teachers to incorporate new reading instruction strategies and materials. The study found that teachers tend to incorporate instructional practices that are encouraged by colleagues, while they tend to reject instructional practices that are communicated as mandatory. My study examines the teachers' responses within a strong test-based accountability policy context where improvement of student academic performance on the test is viewed as mandatory.

Second, my study examines how organizational context shapes teachers' responses to institutional pressures. Existing studies find that the contexts in which teachers encounter policy pressures influence their responses (Coburn, 2001, 2004). To understand how organizational context shapes teachers' responses, I collected data from teachers at two elementary schools that had different organizational structures, different levels of student academic performance, and different levels of available resources.

As it is, my study is in line with recent inquiry about individual responses to institutional pressures. The study adds empirical evidence to better understand the process through which institutional pressures shape individual behaviors, as well as the role of organizational context in this process.

3.4 COGNITIVE APPROACH TO POLICY IMPLEMENTATION

New Institutional Theory has recently begun to examine how individuals respond to institutional pressures. Although institutional theorists have criticized their neglect of agency in examining organizations' and individuals' response to institutional pressures, they have not yet developed a comprehensive framework to explain why individuals respond differently to institutional pressures. Drawing on Giddens' (1984, 1979) concept of structuration, they propose the "duality" of the institution; that is, they conceptualize the institution as both constraining and enabling individual actions. Structuration theory views individuals as knowledgeable agents who interpret, select, and adapt institutional pressures as they take action, but institutional pressures also condition their capacity to interpret and take actions (Scott, 2001). Structuration Theory suggests a relationship between institutional pressure and an individual's response; however, as interaction and knowledge play roles in their responses, Structuration Theory does not provide ways to analyze how an individual's knowledge shapes response. Hence, drawing on Coburn (2004, 2001), I employ cognitive approaches to implementation study (Spillane, Reiser &

Reimer, 2002; Weick, 1995) to conceptualize why individuals respond to institutional pressures so differently.

The cognitive approach is based on the assumption that teachers respond to policy pressures differently because they notice and interpret them differently. It suggests that individuals' knowledge, experience, and context shape their behavior. I draw on this conceptualization to investigate how teachers change their classroom instruction.

Cognitive Approach presents the following three factors as shaping teachers' interpretations of policy pressures: "a) existing knowledge (i.e., schemata, expertise and past experiences), b) their situation, and c) policy signals" (Spillane, Reiser & Reimer, 2002, p. 388). Cognitive Approach suggests that individuals may not consciously notice all information from the policy pressure, but rather information noticed by the individual is *unconsciously* selected and interpreted. Thus, the Cognitive Approach assumes that an individual's sense making is "fraught with ambiguity and difficulties" (Spillane, Reiser & Reimer, 2002, p. 391). Below I present how Cognitive Approach views the role of: a) existing knowledge, b) situation, and c) policy signals, in individuals' interpretations of policy pressures.

Existing knowledge is viewed as an interpretive framework, held by individuals, to filter policy signals. Individuals' existing knowledge varies because it is shaped by past experiences, values, schema, beliefs, and worldviews. A schema is an individual's understanding of causal relationships between things, which is shaped by their experiences and values. Schema assists individuals in making sense of policy pressures by filling in information gaps (Spillane, Reiser & Reimer, 2002). This process of filtering is biased because individuals tend to register messages that they expect and that

are in line with their existing knowledge. An individual's sense-making is a conservative process because individuals tend to interpret information that carries different conceptions as a variation of what they already know (Spillane, Reiser & Reimer, 2002). Individuals tend to pick up superficial information rather than understanding a conceptual difference between the information and their existing understanding.

This conceptualization explains why teachers implement new instructional activities with old and familiar pedagogical strategies (Spillane, Reiser & Reimer, 2002; Cohen, 1990). It also explains why teachers who receive the same policy messages and resources, such as attending system-wide professional development sessions and receiving textbooks, implement policy pressure differently.

Cognitive Approach emphasizes the importance of situations in shaping an individual's sense-making. Situations limit the cultural references people can make. Situations also limit or facilitate access to information. Social values and conceptions about the roles of schools, and ideas about good instruction shared in professional communities and schools, shape teachers' existing knowledge by producing taken for granted assumptions and cultural frameworks. Situations also facilitate teachers' interpretations by providing resources—professional networks and interactions—for accessing new information, which could reinforce or call into question individuals' sense-making and existing knowledge. Immediate contexts, such as instructional activities, also shape teachers' sense-making. For example, a study of implementation of a science curriculum found that some instructional activities facilitate interpretations that are intended by instructional policy while other activities do not (Spillane, Reiser & Reimer,

2002). As it is, advocates of the Cognitive Approach view social context, which includes networks, materials, and activities, as both enabling and bounding individuals' interpretations.

Researchers who employed Cognitive Approach found that forms of representation of policy messages, carriers of the policy messages, and frequency of encounters with policy pressure all influence an individual's interpretation. As discussed in the section about pre-existing knowledge, individuals tend to notice messages that match their prior experiences; hence, policy messages that are in line with their existing knowledge are more likely to reach teachers (Spillane & Jennings, 1997). Carriers of policy messages also influence individual interpretations. People tend to notice information carried by people or sources they trust and respect (Starbuck & Milliken, 1988). Frequency of encountering policy messages also influences teachers' sense-making (Coburn, 2001, 2004).

Summing up, Cognitive Approach offers ways to examine how institutional pressures influence individuals' responses. By examining the above three dimensions that shape individual teachers' responses to policy pressures and how individual teachers interpret policy pressures, the study describes in what context and what aspects of high stakes testing policy pressures influence teachers' instruction. Below I present how I conceptualize my study by using concepts from both New Institutional Theory and Cognitive Approach.

3.5 CONCEPTUALIZATION OF THIS STUDY

This is a study of an organizational field of schools in the context of one district. The study examines how institutional logics changed as the district introduced new standards and assessment in response to NCLB and how these institutional logics are carried to teachers. I view the District of Columbia Public Schools (DCPS) as an organizational field. Bounding DCPS as an organizational field is reasonable for my study because DCPS created its own standards and assessment in response to NCLB. According to New Institutional Theory, I define major actors in the organizational field as education providers, including district government, intermediary organizations, and non-profit organizations that provided services for creating and implementing standards and evaluations[5].

The diagram below represents the conceptualization of this study. The study compares two time periods, prior to NCLB and after NCLB, and examines the change in institutional logics, the spread of these logics, and how teachers changed their instruction. The institutional logic for this study includes the following aspects.

- Concept of how to improve student academic performance
- Concept of how assessment should be used

[5] There is a separate board to govern charter schools in DC. Charter schools follow DCPS standards. Besides meeting with charter school board standards for operating schools, charter schools do not propose different types of instructional policies to influence DCPS's policy or education community. As I describe later, I found most of the instructional messages originate from the DCPS central office and intermediary education providers. I did not find instructional messages from the charter school board. Although both charter schools and regular public schools use the same standards, they do not share the same organizational field. Hence, I view the organizational field for this study as DCPS.

- Concept of good instruction to improve student academic performance. This concept includes the following aspects:
    - How students should learn academic subjects
    - Role of students and teachers in teaching and learning of the academic subjects

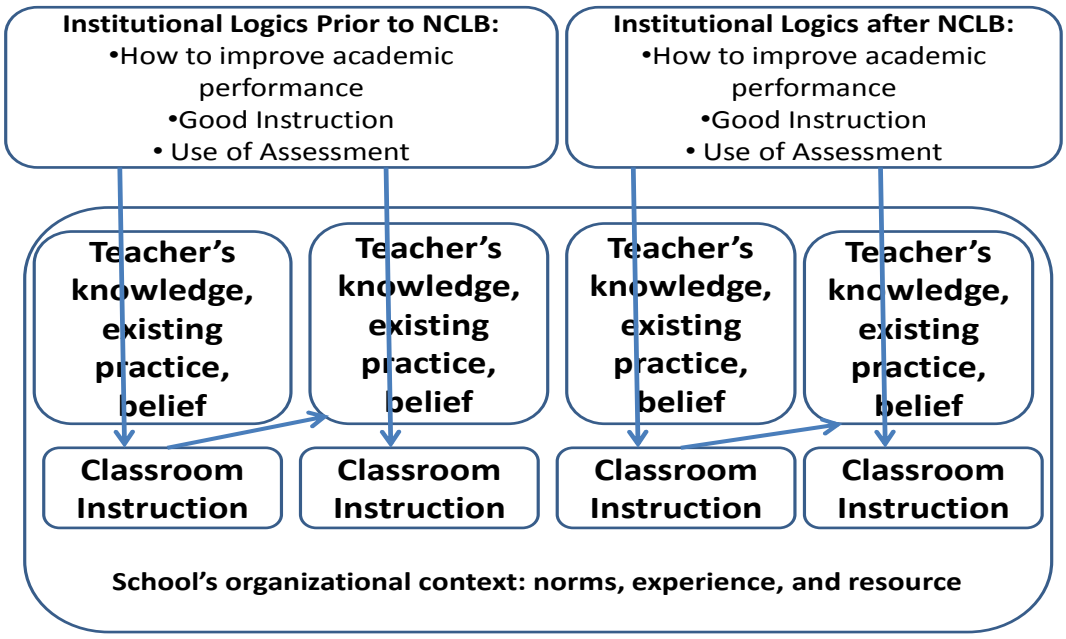

Figure 1 Conceptual Framework of This Study

Drawing upon the studies on the change of U.S. education policy and assessment I included the concept of how to improve student academic performance to identify institutional logics as compared with national level logics. Since the 1980s, U.S. education policy has been focused on how to improve student academic performance (Carnoy, Elmore & Siskin, 2003; West & Peterson, 2003). At the national level, the logic underlying instructional policies shifted from the idea that improving academic

performance means making teachers teach *for* the standards to improving academic performance by using assessment to hold schools accountable for that performance. Policies in the late 1980s and 1990s focused on establishing and aligning learning standards, state assessment, and professional development based on the assumption that this comprehensive approach would make teachers understand standards and that teachers would teach to these standards once they knew them. The federal NCLB Act from 2001 is based on the assumption that teachers improve instruction when their school is made accountable for test results. By reviewing instructional policies issued by the district and other education organizations in the field of DCPS, I identified the conceptual change envisioned by teachers about appropriate strategies to improve student academic performance.

I also examined the use of assessment. Drawing upon studies of change in assessment policy and studies on the impact of high stakes testing (Watanabe, 2007), I examine how the conceptualization of the role of assessment has changed as the DCPS introduced assessment that was aligned with the standards.

I also examined the district policy's conceptualization of good instruction. The existing studies of education policy identified the changes in instructional policy around notions of subject knowledge, how students should learn, and the roles of students and teachers (Cohen, 1990; National Council of Teachers of Mathematics, 2000). Hence, I examine how students should learn math and science according to the district policy and the roles of teachers and students in teaching and learning to identify underlying ideas about good instruction. I compared the change of institutional logic and instructional

change in math and science to understand how high stakes and low stakes testing shape teachers‘ instructional changes.

The study identifies the change of DCPS‘s organizational field from 1989 to 2008. DCPS introduced new math standards and assessment during the school year 2005-06 and introduced new science standards and assessment in the 2006-07 school year to comply with NCLB. As I reviewed the change of instruction and assessment policy in DCPS, I found the institutional logics associated with instruction and assessments prior to the school year 2005-2006 were not different from the logics in the late 80s. In addition, DCPS administered high stakes testing in the late 90s. Hence, I present the change in organizational field of DCPS from the late 80s.

Institutional pressures influence teachers in different ways (Coburn, 2001, 2004). Drawing from Cognitive Approach, I assume that the nature of policy pressures, context, and teachers‘ existing knowledge shape how they respond to policy pressures. I examine an encounter with policy pressure as the unit of analysis to examine teachers‘ responses.

I define policy pressure as any materials and activities that contain messages about instruction and assessment. Policy pressures come from both system and non-system actors, and they take various forms, from formal policy to suggestions about instruction given by colleagues. Defining policy pressure in this broad way (i.e. pressures include both official and non-official policy, including artifacts and activities that convey messages about assessment and instruction) is beneficial for this study because it allows me to document anything teachers think shapes their instruction. In line with other studies (Coburn, 2004, 2001; Spillane & Jennings, 1997), the teachers in this study rarely encountered a complete policy; rather they encountered only pieces of

policy. For example, teachers often picked up tips of instruction as they attended professional development sessions..

To document teachers‘ responses to policy pressures, I employed Coburn‘s (2004) categorization of instructional change because it captures a wide range of instructional change. Instructional changes are: rejection, symbolic response, parallel structure, assimilation, and accommodation. For example, while parallel structure means the new instructional practices do not affect a teacher‘s views of good instruction and student learning, accommodation means the instructional change did, in fact, change teachers‘ beliefs about instruction. Coburn‘s (2004) categorization can be understood as follows:

a) Rejection: Teachers do not take any aspect of the message into their instruction. Teachers either explicitly reject or ignore the message. There is no change in content and pedagogy in instruction.

b) Decoupling/Symbolic Response: Teachers take policy messages and introduce them in their classroom, but they do not change their instruction. An example is hanging state standards in the classroom, but not changing content and pedagogy to reflect them. This incorporation is aimed at showing outsiders that the policy message is incorporated, but it is not aimed at changing students‘ learning.

c) Parallel Structure: Teachers incorporate policy messages into their instruction by creating two or more different approaches in their instruction. For example, teachers teach for the test by using test preparation worksheets for some time, but they also continue instruction that does not incorporate ideas or approaches used in test preparation.

d) Assimilation: Teachers incorporate policy messages into their instruction, but they incorporate it in ways that are in line with their pre-existing practices and knowledge. ―Teachers understand messages in terms of changes in instructional routines, materials, or classroom organization, rather than in terms of the underlying pedagogical or epistemological assumptions about the approach" (Coburn, 2004, pp. 224-225). Often teachers do not notice that their incorporation of policy message is not in line with the intention of policy (Cohen, 1990). Content and pedagogy change but the change will reflect teachers' pre-existing views of student learning and subject knowledge.

e) Accommodation: Teachers incorporate policy messages on the level of changing their pre-existing knowledge. Instructional changes include change in fundamental pedagogical strategy or change of pre-existing knowledge, including conceptions of student learning, subject knowledge, or values.

This categorization positions the change in instruction in relation to teachers' existing beliefs and practices. Hence, instructional practices could be different even though the responses are categorized as the same. Throughout my findings section, I provide a description of instruction.

4.0 RESEARCH QUESTIONS

Drawing on New Institutional Theory and Cognitive Approach to Implementation Study, the research questions for my dissertation study are:

3. How did the institutional logics change as DCPS initiated new instructional and assessment policies in response to NCLB?
    a) Who are the major actors who promoted new policies and activities related to testing and instruction?
    b) How did the institutional logics change along the following dimensions?
        - Conceptions of how to improve student academic performance
        - Conceptions of good instruction

            (Conceptions of role of teachers and students; Conceptions of how students learn)
        - Conceptions of how assessment should be used
4. How did teachers change their classroom instruction?
    a) How did teachers encounter policy pressures associated with instruction and assessment?

b) How did policy pressures become evident in their classroom instruction?

5.0 METHODS

5.1 DESIGN OF THE STUDY

As the goal of the study is to describe the process through which high stakes testing influences teachers‘ classroom instruction, I employ qualitative data collection and analysis. The ethnographic nature of data collection and qualitative data analysis is essential for Institutional Theory to ―explain how social actors and the patterns of action they engage in are institutionally anchored" (Powell, 1991, p. 201). Qualitative data analysis can provide descriptions of how institutional pressures and contextual factors shape teachers‘ classroom instruction.

I collected data from one school district, District of Columbia Public Schools (DCPS)[6], which implemented new district standards and aligned assessment in the school year 2005-06 to fully comply with the requirement from NCLB. DCPS is a single school district in the District of Columbia. Although DCPS administered district tests prior to NCLB, and the results of the tests were used to evaluate principals and to decide whether schools should be restructured, the assessment was not aligned with the state

[6] DCPS functioned as a state as well as a district until 2007. In the District of Columbia there were three entities that govern schools, a) DCPS, b) Charter School Board, and c) Board of Education for charter schools. In 2007, the District shifted the function of overseeing federal grants, state assessment, standards and to the State Office. Until then, DCPS was responsible for creating standards, assessments. .

standards. To comply with NCLB, the state created new state tests that align with the standards, called the District of Columbia Comprehensive Assessment System (DC CAS). Based on this policy context, I collected data to examine how institutional logics changed as the district implemented both the new standards and DC CAS in 2005-06, and how teachers responded to the policy pressures associated with these changes.

This study uses multi-level and cross-sectional qualitative data collection and analysis. I collected various forms of policy documents from 1989 to 2007 in order to understand how institutional logics have changed as the district introduced new standards and assessment in 2005-06. The study examines the logics since 1989 because this is the year the district began to promote new approaches to instruction, and this promotion of new instruction lasted until the new standards were implemented in 2005. I asked teachers about the change of instruction since they had began teaching until the time of the interview, during the school year 2007-2008, to capture the instructional change over time. The study is multi-level because I examine the institutional logics promoted in the organizational field and their implementation at the classroom level.

The study is cross-sectional because I examine the change of instruction of math and science teachers at two elementary schools. In DCPS, math was part of high stakes testing to measure a school‘s Adequate Yearly Progress (AYP), while science testing was not counted in the AYP calculation. This cross-sectional design allows me to investigate how high and low stakes tests differently influence teachers‘ instruction. I selected teachers in this study based on a theoretical sampling to acquire maximum variance in schools‘ organizational contexts and how the nature of the stakes shapes teachers‘

responses. Below I describe how I collected and analyzed the data according to the research questions.

5.2 DATA COLLECTION AND ANALYSIS FOR RESEARCH QUESTION 1: CHANGE OF INSTITUTIONAL LOGICS

To identify the change of institutional logics, I reviewed policy documents, math and science standards, and professional development materials issued by the district from 1989 to 2008. I also interviewed district officials, school level subject specialists, and personnel at a non-profit organization that was involved in creating the district standards to understand the background of these policies. I reviewed newspaper articles from the *Washington Post* and *Education Week* from 1989 to 2008 and evaluation reports created by external organizations to identify major actors and plans for implementation of math and science instructional initiatives, and the results of implementation of the initiatives.

I listed events, such as meetings, publication of strategic plans and standards, and professional development sessions associated with math and science instruction and assessment from 1989 to 2008 based on my interviews and reviews of newspapers, and policy documents. I took note of each standard, strategic plan, and professional development material about how proposed initiatives envision strategies to improve student academic performance, good instruction, and use of assessment. I also looked at the plans for the implementation of these initiatives and compared them with later

evaluation reports to identify actual implementation, revealing the extent to which the instructional approaches and ideas promoted by the initiatives spread in DCPS.

I followed the approach taken by other New Institutional theorists to identify the change of the organizational field (Coburn, 2001; Scott, Ruff, Mendel & Carol, 2000): focusing on changes of governance and institutional logics. In the case of DC's public schools, the governance shaped the implementation of the district policy (i.e. the spread of logics). I identified three eras: a) from 1989 to 1996, b) from 1996 to 2004 and c) from 2004 to 2007, each of which had different governance structures and/or institutional logics about good instruction and use of assessment.

5.3 DATA COLLECTION AND ANALYSIS FOR RESEARCH QUESTION 2: CHANGE OF INSTRUCTION

5.3.1 Data Collection

In order to understand how teachers changed instruction in response to the policy pressures associated with instruction and assessment, I collected qualitative data from eight elementary school math and science teachers at two schools during the 2007-08 school year. I selected math and science instruction to understand if and how high stakes and low stakes testing shape teachers' instructional change in different ways. I conducted a series of interviews with each teacher by using semi-structured, open-ended questions to understand how they encountered policy pressures and how their instruction might

have changed since they began teaching. I also observed their classroom instruction (See Appendix for observation and interview protocols).

Gaining access to the schools and the teachers took a long time. Since my family moved to DC, I wanted to collect data from schools in the DC metro area. I began conducting informational interviews with people who had knowledge and/or experience conducting research in the DC metro area schools. I traveled from Pittsburgh to DC several times between January 2007 and June 2007 and met with educational researchers and district officials in the DC metro area to educate myself on the research application procedure for conducting research in area schools. It was difficult to identify the office that was responsible for research approval in DCPS because the district's organizational structure was changing, and there was a high turnover rate of district officials. After I moved to DC in July 2007, I met with more people, including area non-profit organization leaders, teachers, parents who were actively involved in schools, staff members at DCPS, the Office of State Superintendent, and DC Public Charter School Board. Through talking with these people, I learned how I should present my study, how I should follow-up with people, and who may be able to direct me to the office that approves research requests. In October 2007, I identified the office that was tasked to approve research applications. The research office had moved to the Office of State Superintendent. I met with research office staff members. The research office members were supportive of my study because the office thought an implementation study was needed. The research office staff members suggested a couple of schools that met my first criteria, i.e., one school that had been making AYP and one school that had not been making AYP. The district staff suggested these schools because they knew that the

teachers at these schools were highly committed to their work; hence, the teachers would be likely to talk with me. The research office contacted the principals at the two schools that I call Jefferson Elementary School and Central Elementary School[7] in this study, and I followed up with the principals via e-mail and phone calls.

Access to the teachers at Jefferson Elementary School and Central Elementary School rolled out differently. The principal at Jefferson Elementary School invited me to come to a staff meeting to introduce myself, and she gave me a contact list of teachers after the principal received approval from the teachers. I followed up with the teachers via e-mail and explained what types of teachers I was looking for, i.e., teachers who had taught testing grade prior and after DC CAS was implemented. A total five teachers agreed to be a part of my study. As for Central Elementary School, it took a longer time to meet with the principal. After several follow-up phone calls to the secretary, I finally met with the principal to explain my study. The principal suggested two teachers who met my criteria. The principal talked to each of the teachers, and I followed up. After I began to interview these two teachers, I learned that science was taught by a different teacher. I requested from the principal that I talk with the science teacher, and the principal agreed, as long as the science teacher agreed. All eight teachers I interviewed at these two schools were willing to talk about their instructional practices and their thoughts on instruction, standards, assessment, and other education issues. Thus, my findings are applicable only for these eight teachers and they do not represent all teachers in DCPS.

---

[7] Pseudonym.

The two elementary schools (pre K-6) Jefferson Elementary School and Central Elementary School were different in many aspects. , At the time of selection, Jefferson Elementary School had been continuously making AYP since NCLB was implemented, while Central Elementary School was in year one of improvement status during school year 2007-2008. As I mentioned previously, I originally began my search for schools that had different AYP status in DCPS with the support from the state's Research and Evaluation Office; however, the schools with different AYP status also differed in students' socio-economic status. About 80 % of the students at Central Elementary School were receiving free or reduced-priced lunch, while less than 3 % of students at Jefferson Elementary School were eligible. Students at Jefferson Elementary School were predominantly white, and almost all students in Central Elementary School were African-American. Although the two schools differed in racial make-up, my study did not examine this difference because this study focuses on teachers' instructional change in relation to policy messages associated with high stakes testing. Teaching assignments at these two schools differed. While at M, 5th grade teachers took a "team teaching" approach in which one teacher taught a specific subject, i.e., math, reading and science, for all 5th grade students by students rotating classrooms, at Central Elementary School each classroom teacher taught all academic subjects, including one period of science. There was, however, also a science specialist teacher who was responsible for teaching science to all students at U. Hence, at U, students were taught science from two teachers, i.e. by a science specialist teacher and a classroom teacher. In addition, at Jefferson Elementary School, more teachers were willing to participate in my study, while at U, I could identify only two classroom teachers who had been teaching at Central Elementary

School prior to the new standards and assessment were implemented and who were willing to participate in my study, besides a science specialist teacher. I interviewed five teachers at Jefferson Elementary School, and I interviewed three teachers at school Central Elementary School. The teachers‘ years of experience range from four years to more than twenty years.

Table 1 School Demographics

School	Number of students	Number of teaching staff	Number of low income
Jefferson Elementary School	600 Asian: 4%; Black: 21% Hispanic:4.% Native American 0% White: 71%	34 Student and teacher ratio 17:1	Approx.3 %
Central Elementary School	380 Asian: 0% Black:99%%; Hispanic:1%; Native American: 0% White 0%	38 Student and teacher ratio 13:1	Approx. 90%

Note: All population numbers are from the 2006-2007 school year.I created this table based on the information on DCPS web site.

Table 2 Historical AYP Report

	2003	2004	2005	2006	2007	2008
Jefferson Elementary School (reading)	Yes	Yes	Yes	Yes	Yes	Yes
Jefferson Elementary School (math)	Yes	Yes	Yes	Yes	Yes	Yes

Central Elementary School (reading)	Yes	No	Yes	No	No	No
Central Elementary School (math)	Yes	No	Yes	No	No	Yes *

Note: For school Jefferson Elementary School, Black, non-Hispanic, White non-Hispanic are reported as ethnicity sub-category. As for 2005, black is not reported because this group is less than 25 students. DCPS determines elementary schools' AYP status based on testing participation rate, attendance rate, and proficiency rates on Math and Reading tests.

Note*: Central Elementary School made AYP in math in 2008 by meeting safe harbor requirements (reducing the number of non-proficient students by 10%). Sub group is reported for Black, non-Hispanic, and economically disadvantaged.

I interviewed teachers during their lunchtime or after work, depending on their availability. Each interview session lasted from less than five minutes to over one hour, depending on the schedule of the teachers. I also had informal conversations with the teachers so that the interviews wouldn't be a burden for them. Total interview time for each teacher ranged from one hour and ten minutes to more than four hours. Questions focused on understanding how teachers encountered policy pressures and how their instruction changed over time (see the appendix for interview protocol). I used a digital recorder most of the time. I did not record informal conversations. One time my digital recorder's battery ran out during a formal interview. I transcribed interviews by myself for analysis.

I also observed lessons at least once for each academic subject during the school year. The purpose of this observation was to verify my understanding of their instructional approaches. I took observational notes with a descriptive system and used functional and IRE (initiative, response, and evaluation) units (Evertson & Green, 1986).

Rather than using a pre-made checklist of teacher behavior, I took fieldnotes to document tasks and activities the teacher facilitated, responses by the students, and teacher feedback in a low inferential manner. Documenting activities around these observation units allowed me to capture detailed contexts and the nature of the interactions so that I could characterize the nature of the questioning, assistance, and mathematical tasks the teacher gave to the students. These characteristics are important to identify the nature of instruction, particularly when the instruction is characterized as reform-oriented and traditional.

Table 3 Data Collection from Teachers

Teacher	Years of experience (year in DCPS)	School	Academic Subject	Grade	Number of interview session	Total duration of interview	Number of observation	Total duration of observation
Mrs. Reed	18	Jefferson	Math	4th	5 times	1: 11:13	3 times	2:45
Mr. Shawver	4	Jefferson	Science	5th	3 times	1:20:02	2 times	2:30
Mrs. Canright	5	Jefferson	Math	5th	5 times	2:12:02	3 times	2:45
Mrs. Caracelli	8	Jefferson	Math	3rd	4 times	1:22:51 (approx)*	2 times	1:40
Mrs. Miller	6 (5)	Jefferson	Math and science	4th	8 times	1:59:41	Math 3 times	Math 2:45
Mrs. Green	4	Central	Math and science	5th	9 times	2:20:54 **(approx)	Math 1 time Science 6 times	Math 2:30 Science 10:30
Mrs. Sommers	32 (4)	Central	Math	4th	11 times	4:27:16 (approx) ***	Math5times	Math 6:50 Reading
Mr. Berry	15	Central	Science	K-6	3 times	2:02:46	5 times	3:30

*: Mrs. Caracelli: one time the digital recorder went off, so I have estimated the duration of the interview
** Mrs. Green: includes a note of conversation.
*** Mrs. Sommers: includes four notes of conversation (total approximately Oue hour and 15 minutes) and one time the digital recorder went off, so I estimated the duration of the interview.

5.3.2 Data Analysis

I analyzed qualitative data I collected through the interviews and observation as follows. After each observation, I read my fieldnotes and summarized the nature of instruction around the following aspects: a) nature of tasks, including size and complexity of problems teachers assigned; b) students‘ activity and discussion on problem solving; and c) classroom interaction, including whether students worked with others to discuss math and science concepts and problem solving strategies. I created a list of the nature of tasks, processes of student exploration, and patterns of interaction based on DC‘s academic standards and created a description of how particular lessons were in line with the approaches promoted by the standards. For example, the math standards indicate that a teacher should encourage students to explain their problem solving procedures because students should ~~–e~~larify mathematical ideas as they discuss them with peers and reflect on strategies and solutions” (District of Columbia Public Schools, 2005, p.4). This instructional approach suggested in the math standards was in line with the recommendations of the National Council of Teachers of Mathematics (NCTM, 2000), and the district‘s strategic plans often emphasized moving instruction from a traditional teacher-centered approach to a student-centered instructional approach. After I read the

fieldnotes of each lesson, I wrote down if and how a teacher encouraged students to do so in a particular lesson I observed. Writing down these observations proved valuable, creating a recorded and explicit description of instruction in order to compare how instruction differed among teachers.

I compared my observation notes with teachers‘ interview responses, particularly with their descriptions of good instruction and how students learn. This comparison was helpful in understanding the meanings of the terms teachers used to describe their teaching practices. For example, teachers often said that it is very important that students understand the mathematical concept behind a mathematical operation. The observation informed me that teachers worked differently in terms of how each facilitated tasks aimed at furthering students‘ understanding of concepts. Thus, the nature of interaction and the range of students‘ exploration varied from teacher to teacher. One teacher presented the mathematical tasks from *Everyday Mathematics* to explore the concept of ratio and percentage and let students work on the tasks individually, while another teacher welcomed students‘ comments about how they imagined the connections between the concepts and their other experiences. Both approaches to instruction were aimed at providing a conceptual understanding, but the range of exploration students experienced and interaction among students and teacher were quite different. I asked the teachers follow-up questions based on the observations to understand whether the activities I observed were typical for them and why they did the activities in a particular way.

I coded transcribed interviews by using codes I made from the theoretical framework. These codes included the nature of policy pressures and types of responses. They are derived from the Cognitive Approach to Implementation as the factors that

shape individual teachers‘ different responses to policy pressures. I created sub-codes under pre-made codes to manage the volume of the data, as well as to articulate the differences under pre-made codes. For example, teachers provided descriptions about how testing influences their instruction, and some descriptions touched upon their views of how they should prepare students for the test. Although this description could be merged into the code about their views of how students should learn mathematics and science, I coded the description separately to understand how teachers see the specific requirements from the test. I coded them using the pre-made codes of ―policy pressure" and ―response." I also utilized open coding to capture the variety of instructional approaches teachers take. Under the pre-made code of ―response," I created a sub-code ―preparing for tests" and then created sub-nodes that labeled different types of test preparation the teachers take. At the same time, I assigned a conceptual code. I drew from Coburn‘s (2005) categorization of teachers‘ instructional responses to the texts coded under descriptive codes p. As it is, I coded data under general descriptive and conceptual codes that derive from the theoretical framework. Then I added both conceptual codes and open codes to capture the variety of activities and views by using open coding.

I constantly compared the texts under each code within and between teachers and asked follow-up questions in the interviews. For example, when one teacher mentioned the importance of making students complete benchmark tests within a given time, I asked other teachers whether they see this as an important part of their teaching for the test. I also crosschecked the implications of these answers to the conceptual codes.

I used QSR NVivo 7 to assist with data management and retrieval. I assigned academic subject, years of experience, and school to each document to create cases. Having cases helped me to easily retrieve data that are associated with dimensions I anticipated to shape teachers' classroom instruction. For example, I could generate a code report for all math teachers or all teachers from a specific school.

After I coded the interview transcripts, I read the texts in each code and wrote up an analysis summary for each teacher. The analysis summary included the teachers' views of good instruction, their sense of how students learn, and their understanding of the nature of math and science knowledge, taking into account a comparison of the interviews with the analysis of the observation. I used the search function of NVivo 7 to assist with constant comparison. For example, I retrieved a node report of "parallel structure" for each academic subject to see if math teachers tended to respond in a specific way more than science teachers. In the summary, I also listed policy pressures the teachers came across and their responses. I later categorized policy pressures by the nature of policy messages, i.e., if the teacher included instructional message or not and if policy pressures included monitoring. I followed the codes of responses from Coburn's (2005) categorization and summarized how each teacher responded to each policy pressure, finding patterns within one teacher according to the nature of policy pressures they came across.

Then I compared the nature of policy pressures and teachers' responses between math and science, and then between schools in order to understand how the nature of policy pressures associated with assessment and organizational context shaped their instructional change.

5.4 GENERALIZABILITY, VALIDITY AND LIMITATIONS OF THE STUDY

5.4.1 Generalizability

Generalizability relevant to this study can be understood as the following two kinds: ―user generalizability" (Merriam, 2002, p. 28) and ―vertical generalizability" (Silverman, 1993, p. 180). The goal of qualitative research is providing ―in depth understandings and knowledge of particular phenomena" (Schofield, 2002, p. 100), rather than showing the effect of particular interventions on a general population. Hence, the findings from my study will not be extrapolated to all schools and teachers. The findings are from the eight teachers at the two specific schools in which I collected data. Although my analysis is based on the limited number of data sources, the findings could inform readers and researchers about how high stakes testing shapes teachers' instruction and how different factors shape this process so that they can transfer concepts from my study to other contexts with the specific context of my study in mind (user generalizability).

―Vertical generalizability" (Silverman, 1993, p. 180) refers to the extent to which my particular study contributes to theory. I designed my study by drawing upon New Institutional Theory and a field of study on teachers' instructional change in response to high stakes testing. I presented how I defined the concepts available from existing theoretical and empirical studies on these two topics in the conceptual framework section. As I will discuss in the findings sections, my study will elaborate on and expand existing theory by presenting empirical evidence that delineates concepts used in New Institutional Theory. My study also presents findings that were less explored in existing

studies on teachers‘ instructional change in response to high stakes testing. In this sense, my study is an attempt for ―theoretical elaboration‖ (Vaughan, 1992). I present the process of theoretical elaboration, i.e., how I used concepts from New Institutional Theory and Cognitive Approach, and how my data delineated some concepts.

5.4.2 Validity

Validity asks, ―how congruent are one‘s findings with reality‖ (Merriam, 2002, p. 25). Drawing on Maxwell‘s (2002) typology of validity in qualitative research, I describe how I constrain the threat to three types of validity, i.e., descriptive validity, interpretive validity, and theoretical validity.

Descriptive validity refers to the factual accuracy of the accounts of my respondents (Maxwell, 2002). I used a digital recorder as much as possible. I took fieldnotes with non-inferential descriptions. In order to increase the accuracy of my findings, I consulted different data sources to triangulate the information, for example, between interview and the district‘s web site. I used an iterative and reflexive process by going back and forth between interview and document review to clarify my understanding of events and activity. It was challenging to acquire materials associated with standards and professional development, as the district did not have a centralized archive system. To fill in the gap of information, I used newspaper articles and reports created by non-profit organizations as well as informational interviews with participants in professional development to check the nature of the event associated with the

standards and assessment. I interviewed teachers multiple times to increase the accuracy of description the teacher provided to me.

Interpretive validity refers to the accuracy about what objects, events, and behaviors mean to the people engaged with them (Maxwell, 2002). In order to have a rich and clear description of a teacher's views of instruction and the change of instruction, I employed the following three strategies. First, I coded teachers' descriptions of instructional change by using open coding and by using a teacher's terms once I coded them under a broad descriptive code, such as "response." Open coding helped me to consider a teacher's interpretation of the situation, rather than assigning pre-determined meaning to a particular teacher's account. For example, teachers provided various forms of test preparation. Before I coded them into a pre-made code of "parallel structure" or "assimilation," I coded them by using their terms so that I could acquire more detailed descriptions about the activities and intentions teachers described. Second, I made a case report for each teacher about how they teach by using their terms with clear definitions. Then, I compared the terms with the observation to develop a more concrete image of each teacher's instruction.

Theoretical validity refers to an account's validity as a theory of some phenomenon (Maxwell, 2002). The threat to this validity is that I might build theories using selective evidences. To minimize this threat, I analyzed data by using the following techniques. First, I constantly compared data and concepts. Additionally, I created a series of statements pertaining to patterns of factors that shape specific teachers' responses, and I checked with the data to clarify the extent to which specific statements were valid. I also looked for counter evidence (Silverman, 1993) and any third variables

that might be shaping teachers‘ instructional change by asking ―why‖ specific factors seem to be shaping teachers‘ instructional change (Eisenhardt, 2002).

5.4.3 Limitations of the Study

Limitations of this study come from the data collection methods. First, this study relied heavily on interviews to identify a teacher‘s instructional change. Interviewees might have re-conceptualized past experiences and thoughts, which may not reflect the actual processes of their thinking and their actual instructional practices. In order to increase the validity of interviewees‘ accounts, I constantly compared the factors they mentioned, asked for examples, and looked for other sources. For example, the teachers did not have clear memories about professional development sessions that happened within schools. In this case, I asked several teachers about it to validate the information.

Second, I might not have found all the factors that shape instruction as well as all the instructional changes teachers made or all the policy pressures they might have encountered through the interviews and the limited number of observations. To minimize this limitation, I reviewed the literature on how teachers shape instruction and asked whether similar factors influence their instruction. I also reviewed policy documents and newspaper articles and listed policy pressures the teachers might have experienced.

Third, my presence might have influenced teachers‘ classroom practices. For example, when I observed lessons, one teacher asked me how I would make sure the students understand how to simplify fractions. I provided my suggestions, and the teacher tried the suggested approach for thirty minutes. On another occasion, one

teacher introduced me as a researcher from Japan who studies U.S. schools in an attempt to make her students pay attention to a lesson. To understand how my presence might have influenced the teachers‘ instructional practice, I took notes of interaction patterns and teachers‘ questioning in these situations and compared them with other situations. The teachers responded that the lessons I observed were typical, and the patterns of questioning were consistent across lessons I observed.

Fourth, the document data I could obtain to capture the change of organizational field was limited. Although I extensively contacted district officials and teachers to ask about instructional materials and PD materials used since 1989, I could not obtain them. The documents, particularly those associated with professional development, I could obtain were mostly from websites or through my interviewees; hence, it was not comprehensive and I likely missed some events and instructional approaches promoted. Furthermore, some of the important information about policy messages was missing from my data set. For example, I did not look at the SAT test and DC CAS tests, which seem to be two of the most powerful carriers of instructional messages for teachers. As a result, I heavily relied on publicly available evaluation reports and strategic plans to infer the instructional practices promoted. The description of the nature of the tests was based on interviews and newspapers.

In addition, there is a possibility that these documents may not reflect cognitions of all actors and activities in the field (Schneiberg & Clements, 2006). For example, I reviewed the Washington Post, a local newspaper, to find out about activities conducted by various groups in relation to math and science instruction and testing; however, there might have been other activities that may have carried ideas about instruction and testing.

As it is, the description of the change of organizational field is based on publicly available documents and the small number of documents I could obtain in hard copy from district officials and non-profit organizations. Throughout the findings section, I present the source of the data to help keep readers informed about the scope of the data I used.

6.0 FINDINGS: CHANGE OF ORGANIZATIONAL FIELD

6.1 INTRODUCTION

In this section, I present how the organizational field of the DC Public Schools changed as the district created policies to respond to the requirements from NCLB. To describe the change of organizational field, I reviewed policy documents, such as strategic plans, guidelines from the district, evaluation reports and newspaper articles to identify what types of good instruction are promoted in these documents and how the use of assessment was conceptualized.

When I began the document review, it became clear that the concept of good instruction promoted by the district's standards had not changed in response to NCLB in 2005. The change NCLB brought to the district was that it aligned assessment and standards. Until then, the district had standards and high stakes testing, SAT9, and they were not aligned. Although the district presented the new standards as drivers to change classroom instruction in 2005, the concept of good instruction promoted in the new standards was the same as the one promoted in the old standards. The district has promoted reform-oriented math and science instruction since 1989.

In this section, I describe the organizational change of DCPS from 1989 to 2008, the year I collected data from teachers. I describe the change of organizational field of

DCPS by focusing on the change of institutional logics and governance structure. In DCPS, the governance structure of the district shifted from site-based management to centralization of instructional policy in 2005. The change of governance changed how the district could carry instructional policy to schools and teachers. In this section, I describe the change of organizational field by dividing my focus into three time frames: a) 1989-1995, b) 1996-2003, and c) 2004-2008. These divisions are informed by the change of institutional logics the district promoted and the change of governance. Between 1989 and 1995, the district promoted reform-oriented instruction through issuing a series of strategic plans. Between 1996 and 2003, the district administered high stakes testing, SAT9, which was not aligned with the standards. During this time period, newspaper articles reported that much instructional time was spent on drilling and worksheets in order to prepare students for the test, which suggests that the district promoted two different ideas of what classroom instruction should look like. One is reform-oriented instruction promoted by the standards, and another is teaching for the test promoted by SAT 9, which the district introduced. Although the district continuously promoted reform-oriented instruction since 1989, in line with the national level standards-based education movement, the district did not have measures to spread instructional policy and guidance until 2005 because the district had allowed schools site-based management. In 2005, the district centralized instructional decision-making and required schools to follow the standards. In 2005 the district introduced a new high stakes test, DC CAS, to comply with the requirement from NCLB.

Thus, in DCPS, the policy change in response to NCLB was different from the national trend. While the focus of education policy at the national level shifted from

establishing academic standards in the 1990s to the emphasis on test-based accountability as NCLB was issued, DCPS promoted a test-based accountability in the late 1990s. NCLB required the DCPS to align its standards with assessment.

To describe the change of organizational field, I offer the following descriptions for each time period: a) institutional logics of good instruction and how assessment should be used and policy pressures that promoted these specific conceptions, b) governance of the district, and c) spread of institutional logics of good instruction and use of testing.

Throughout this section, I describe how different forms of governance and carriers of institutional pressures promoted different conceptualizations of good instruction. The table below presents the change of DCPS' organizational field. Between 1989 and 1996, the district promoted uniform institutional logics about what good instruction should be. The district promoted reform-oriented instruction, departing from traditional, teacher-centered instruction. Between 1997 and 2003, the district then promoted conflicting conceptions of good instruction because policy pressures associated with the standards and high stakes testing, SAT 9, promoted different ideas about how classroom instruction should take place. The standards encouraged reform-oriented instruction, but the newspaper reported the teachers were teaching for the test by using worksheets and drilling. In 2004, the district began to promote a unified message about good instruction by creating new standards and by aligning assessment to the standards. This time, the district required all schools to follow the standards. In 2007, when the district began to provide intensive support to the schools to prepare students for DC CAS,

the policy pressures associated with DC CAS carried different concepts of good instruction.

Table 4 Change of DC Public Schools' Organizational Field

Year	Superintendent	Events	Institutional Logic about good instruction promoted by the district		Governance
1989 - 1996	Andrew Jenkins (-1990) Franklin Smith (1991 -1996)	1989 –Our Children, Our Hope" report published	Reform Oriented Instruction		Site-based management
1996 -2004	Julius Becton (1996- 1998) Arlene Ackerman (1998-2000) Paul Vance (2000-2003) Elfreda Massie (2003-2004)	1996 Congress oversight 1997 High stakes test, SAT 9 started 2000 School Board recovered authority	Reform Oriented Instruction	Teaching for Test	Site-based management
2004-2008	Clifford Janey (2004-2007) Michelle Rhee (2007-2009)	2005-06 Math, Reading Standards DC CAS started 2006-07 Science Standards DC BAS started	Reform Oriented Instruction Teaching for Test		Centralization

6.2 FROM 1989 TO 1995: PROMOTING REFORM-ORIENTED INSTRUCTION

From 1989 to 1995, the district promoted instruction that was in line with the national level standards-based reform movement. The district promoted reform-oriented instruction through which students were to learn math and science by working on problems drawn from life experiences. The district promoted the idea that assessment should be used to improve instruction. The governance system during this time was site-based management, which allowed schools to decide on their own curriculum and instructional approaches.

In this section, I first present institutional logics about instruction and assessment by reviewing instructional policies associated with math and science. The district promoted standards-based reform through the publication of strategic plans. Second, I describe how the governance system employed during this time reflects the district's conceptualization of how to improve academic performance. Third, I discuss that site-based management allowed each school to choose its own instructional approaches; thus, the institutional logics promoted by the district did not appear to have spread to schools.

6.2.1 Institutional Logic: Reform-Oriented Instruction

From 1989 to 1995 the district repeatedly proposed introducing reform-oriented instruction and site-based management as the way to improve student academic

performance. The district issued strategic plans aimed at changing from teacher-centered instruction to student-centered instruction and changing students' experience of learning from rote memorization to conceptual understanding. The district acquired grants from the federal government and from national foundations to promote reform-oriented instruction. The table below shows policy initiatives issued during this time period.

Table 5 Timetable of Policy Initiatives between 1989 and 1996

Year	Policy Initiatives
1989	The D.C. Committee on Public Education published ―Our Children, Our Hope: Revitalizing the District of Columbia Public Schools" (Recommended abolishing Competency Based Curriculum, criticized teacher centered instruction),
1990	District Central office held total 30 meetings to develop strategic plans to implement report's recommendations
1991	―Comprehensive Five Year Plan" by DCPS issued ―Bringing Learning Communities: A Framework for Preparing Students for the 21st Century, The District of Columbia Public Schools Response to America 2000" by DCPS, business leaders, and non-profit organizations issued.
1992	―A Strategic Plan: Five Year Approach" by DCPS issued (support for Local School Restructuring Team)
1994	Teachers attended workshops provided by AAAS's Project 2061 DCPS received a grant from NSF Urban Systemic Initiative Program
1995	―Bringing Educational Service to Students" by DCPS issued
1996	DCPS goes under Congressional oversight.

Evaluation reports, strategic plans, and professional development provided with the funds from national organizations all promoted reform-oriented instruction. This reform began with a 179-page evaluation report, *Our Children, Our Hope: Revitalizing the District of Columbia Public Schools,* published in 1989. This evaluation report, written by a committee of business leaders, community members, and educational

specialists stated that the problem of low academic performance among older students derived from teacher-centered instruction. The report criticized Competency Based Curriculum (CBC), which had been used in the district for the past 15 years. CBC presented critical skills and activities the teachers should make students do based on the progression of student learning, so that students were to be assigned progressively more complicated skills and concepts over time. It presented a plan for each lesson.

In –Our Children, Our Hope," CBC is viewed as the primary reason for teacher - centered instruction because it made teachers focus on checking off skills and activities, which resulted in direct instruction. It stated that although CBC was thought to be the best way to teach students 15 years before, recent research found that students should be exposed to tasks that require both low level and high level thinking. The report states how CBC became the source of teacher-centered instruction and suggests what recent research had identified as better instruction methods as follows:

> CBC was a state-of-art system when it was developed. At that time, there was a widespread belief that children would learn more if learning tasks were broken up into small steps and if students were assessed frequently on their mastery of the ordered sequence of steps. CBC was also based on the belief that more learning occurred when teachers engaged in direct, whole-class instruction. Thus, the curriculum guides of the District contain hundreds of small objectives and teaching suggestions. Over the life of the CBC project, the school systems has attempted to make these cumbersome lists of objectives and teaching suggestions more manageable by providing teachers with abbreviated checklists of items on which students would be tested. Over the past 15 years, though, both research

> and experience have challenged many of the underlying assumptions of CBC. For example, there is now very powerful evidence that children lose the threads of connection between steps in a learning process if the process is broken up into too many small, easily measured objectives. Evidence also exists that teacher-controlled whole-class instruction –although it is appropriate to some learning tasks—does not help children to learn to read for understanding or think critically. (The D.C. Committee on Public Education, 1989, pp. 56-57)

The report argued that providing students with tasks that require higher order thinking from an early age is the key to sustaining academic performance in older grades. The report also stated that a majority of teachers did not use classroom assessments and standardized tests for improving instruction (The D.C. Committee on Public Education, 1989). Following up with this evaluation report, the district published a series of strategic plans that all promoted introducing new curriculum and changing student learning from rote memorization to –learning in order to know" (District of Columbia Public Schools, 1992, p.10).

The district also acquired funding from national organizations to implement reform-oriented instruction. The business community, the district, and community members met and agreed upon the goals for DC schools, which resulted in a report entitled, *DCPS Response to America 2000,* in 1991. A –Five Year Strategic Plan" was published in 1992, which incorporated the action plans developed in *DCPS Response to America 2000*. The plan stated that the district would take systemic reform efforts with the belief that all students can learn to high standards. *DCPS Response to America 2000* also presented specific goals and instructional approaches for math and science. Among

the eight goals, the fourth goal dealt with math and science. *DCPS response to America 2000* states, "Goal 4: Students will be first in the world in science and math achievement," and proposes aligning curriculum in line with the national standards proposed by the National Council of Teachers of Mathematics (District of Columbia Public Schools, 1992, p. 16). These documents suggest that math and science instruction should cover basic skills, critical thinking, higher order thinking, and problem solving. The strategic plans encourage incorporating performance assessment to measure higher order thinking (District of Columbia Public Schools, 1991). Thus, in line with the national reform movement, the idea of good instruction promoted by these strategic plans was teachers facilitating activities that provide students with opportunities to apply knowledge to real life situations, rather than teachers delivering knowledge *to* students. The district policy papers repeatedly stated plans to align standards, curriculum and assessment, and also stated that assessment should be used to improve instruction.

Projects brought by national foundations, such as the National Science Foundation (NSF) and the American Association for the Advancement of Science (AAAS), also promoted reform-oriented instruction. Teachers attended workshops provided by the AAAS' Project 2061 on benchmarks and science literacy (West, 1994, March 23). DCPS also received a grant from the NSF's Urban Systemic Initiatives Program to aid in reforming math and science instruction in November, 1994 (West, 1994, September 14). These funds were intended to help to change classroom instruction to reform-oriented instruction in line with the national trend.

During this period, the major carriers in promoting new instruction seemed to be official policy documents and occasional professional development opportunities. I did

not find any reports that monitored if teachers were actually teaching reform-oriented instruction. Although the strategic plans addressed establishing benchmarks to measure academic improvement among students by using performance tests, I could not find any documents that presented details about assessment. Furthermore, the district continued to use the Comprehensive Test of Basic Skills throughout this time period. The test results were published, but I did not find any document that described the consequences of their publication; hence, the assessment was a low stakes assessment.

During this time, the district did not manage funds well. By 1996, continuous financial and personnel mismanagement placed the district under the oversight of the United States Congress's Financial Control Board. The Board of Education was dismantled, and the district was to report to the Control Board. The Control Board appointed a new superintendent who became responsible for implementing reform strategies to improve the district's financial situation and its student academic performance.

6.2.2 Governance

Site-based management was the governance structure in the district between 1989 and 1995. The district allowed schools to choose curriculum and staffing, and the schools could use some discretionary funds for their own instructional initiatives. The district believed that site-based management would improve student performance.

The *Our Children* report in 1989 proposed expanding each school's decision-making authority. The district proposed shifting the decision-making authority to each

individual school so that ―principals, teachers, students, parents and community members" could make decisions about ―curriculum, budget, staffing and facilities" (District of Columbia Public Schools, 1995, p. 3). This shift was based on the belief ―that those most closely affected by decisions should play a significant role in making those decisions" and ―that educational reform efforts will be most effective and long lasting when carried out by people who feel a sense of ownership and responsibility for the process" (District of Columbia Public Schools, 1995, p.3). The district established an office to promote the Local School Restructuring Team (LSRT) through which parents, principals, and teachers could make decisions about their own schools.

Site-based management was proposed in line with promoting school choice. The report, *Our Children, Our Hope*, states that each school is to become a ―school with distinction" (D.C. Committee on Public Education, 1989, p. 80), a magnet school. The superintendent proposed district-wide open enrollment in 1992 and argued schools should develop them and that all DC schools should be magnet schools by 1996-97. The superintendent also proposed that schools that could not attract students would close (Smith, 1992, July 12). Although complete open enrollment did not happen, students in DC could attend any school if there was an opening at that particular school for out of boundary students. Site-based management became a norm in DC's public schools. As described in an evaluation report published later, schools were left to provide instruction of their own choice (The Council of Great City Schools, 2004).

6.2.3 Spread of Institutional Logics

I could not find any documents that described actual classroom instruction during this time. Evaluation reports and newspaper articles reported that the district had failed to provide the resources to support instruction and failed to follow the recommendations presented by the *Our Children, Our Hope* report. In addition, the district allowed each school to decide curriculum because of the site-based management. Thus, I infer that implementation of the above stated reforms initiative was weak in DCPS, and the concept of reform-oriented instruction might not have reached schools.

Another indication that the reform-oriented instruction was less likely to spread in the district schools was that the district lost funding to promote reform-oriented instruction. (District of Colombia Financial Responsibility and Management Assistance Authority, i.e., Control Board, 1996). The National Science Foundation withdrew funding for the Urban Systemic Initiative Program because of the district‘s financial mismanagement in August 1996 (West, 1996, October 2). The district lost funding from America 2000. Although site-based management was to improve the capacity of each school to plan and select best instruction and resources to improve student academic performance, ―the restructuring teams that are supposed to revolutionize local school management barely exist" (Hening, 1999, p. 112).

It was widely known early on that the district‘s central office had both financial and personnel mismanagement, suggesting a breakdown of carriers of institutional logics associated with instruction and the use of assessment. An auditor for DCPS concluded that the DCPS did not know how many people were on its payroll because accurate

documentation did not exist, and the district did not implement the recommendations of *Our Hope Our Children*. (Education Week, 1993, May 12). The evaluation report, *Children in Crisis*, published in 1996, reports that about 20 percent of schools did not have textbooks in place (District of Columbia Financial Responsibility and Management Assistance Authority, 1996). Taken together, the documents suggest that the reform-oriented instruction did not widely spread to schools.

6.3 FROM 1996 TO 2003: PROMOTION OF REFORM ORIENTED INSTRUCTION AND TEACHING FOR THE TEST

During this time, the district promoted contradicting ideas about good instruction. The district began to use standardized tests to hold schools accountable for student performance. The SAT9 became the measurement of student academic performance, and it became high stakes for schools and principals. At the same time, the district continued to create strategic plans that promoted reform-oriented instruction for math and science. Newspaper articles reported that schools stopped regular instruction and taught for the test.

In this section, first I will present how strategic plans and administering SAT9 promoted different conceptions of good instruction. As I did not have access to actual test problems and I did not find any documents associated with how teachers should teach for the test, my description in this section is based on newspaper articles, strategic plans

created by the district, and evaluation reports published later. Second, I present how site-based management shaped the way institutional logics spread among the schools. Third, I present the spread of institutional logics based on the evaluation reports.

6.3.1 Institutional Logics: Reform-Oriented Instruction and Teaching for the Test

Between 1996 and 2003, the district promoted conflicting conceptions about good instruction. During this time period, assessment became a measure of how well schools were improving students' academic performance. A review of policy documents, evaluation reports, and newspaper articles suggests that the strategic plans and the assessment carried different conceptions about good instruction. The table below lists initiatives and policy documents issued during this time period. I describe how these documents conceptualized good instruction and the use of assessment.

Table 6 Timetable of Policy Initiatives between 1996 and 2004

Year	Policy Initiatives
1996	Control Board issues "Children in Crisis" (low academic performance of DCPS students Superintendent Becton was appointed by the Control Board. Becton begins "Children First" initiative to implement DC Goals 2000.
1997	Chief Academic Officer, Ackerman began to work to improve academic performance. SAT 9 was administered, and it became the baseline to measure improvement of each school. District planned to create school profile The district decided to shorten principal's tenure from three years to one year, and used SA T9 performance as major part of principal evaluation Schools whose performance declined were identified for receiving targeted assistance. These schools later introduced Comprehensive School Reform programs. District began Summer Schools

1998	The district proposed performance pay for teachers
2000	Some teachers participated in creating a draft of academic standards to introduce standards-based national standards. School Board recovered. Superintendent Vance began his term.
2001	―Children First Initiative: Action Plan for Transforming Schools" --focus on improving low performing schools based on SAT 9 results US Department of Education sent a letter to the district informing the district it failed to comply with the requirement for Title I. District agreed to develop assessment that aligned with the standards by the 2004-05 school year.
2002	District Standards, Pacing Chart (Note: Old Standards in this study) **NCLB was signed.** SAT 9 performance became baseline for determining AYP. NAPE administered. DC students performed lowest in large urban districts.
2003	US Department of Education notified the district that the district's accountability plan did not comply with the NCLB requirement. The district must create assessment that is aligned with the standards. Superintendent Vance announced his resignation and Massie was appointed as interim superintendent Vance and Massie requested the Council of Great City Schools to evaluate the district.
2004	―Restoring Excellence to the District of Columbia Public Schools" by Council of Great City Schools

Assessment became an accountability measurement when Congress began to oversee the district. The introduction of a new assessment, SAT 9, was a part of a reform effort to improve DC public schools. Below, I describe the concept of good instruction and the use of tests that are promoted by policy pressures associated with SAT 9.

The Congress took over control of education in DC in 1996 because of the district's continuous budget deficit, financial mismanagement, and fire code violations. The report, *Children in Crisis*, by the District of Columbia Financial Responsibility and Management Assistance Authority (called the Control Board), appointed by the Congress, describes DC's system as ―failing." The report presents low academic performance of students as the indicator of the failure of DC's public schools. This time,

the low academic performance included the early grades up to high school students. The report presents test results of the Comprehensive Test of Basic Skills (CTBS), which DCPS had administered for a decade, and states, "Educational outcomes are inequitable and weak" (District of Columbia Financial Responsibility and Management Assistance Authority 1996, Failure of Public Education in the District section, para 7). Academic performance declined dramatically among students in the poor wards of the city. After showing the test results from 1991 to 1995, the report concludes the failure of DCPS as follows: *"The CTBS scores suggest that the longer students stay in the District 's public school system, the less likely they are to succeed educationally" (*District of Columbia Financial Responsibility and Management Assistance Authority *, 1996*, Failure of Public Education in the District section, para 9, *this emphasis was original)*[8] The report also presented the decline of student academic performance by word in the district as follows.

> Between 1991 and 1996, mathematics scores declined by an average of 6 percent for grades 6, 8, and 11, and reading scores declined by 13.5 percent in grade 6 and remain below national norms in grades 8 and 11. While test scores in wards 1, 3, 4 and 5 have remained the same or slightly improved, scores in wards 2, 7, and 8 have declined significantly. Several schools in wards seven and eight have seen startling declines in test scores of 15 to 20 percentage points or more. Ward 7, for

[8] The report also shows the decline of student academic performance by ward in the district. Between 1991 and 1996, mathematics scores declined by an average of 6 percent for grades 6, 8, and 11, and reading scores declined by 13.5 percent in grade 6 and remain below national norms in grades 8 and 11. While test scores in wards 1, 3, 4 and 5 have remained the same or slightly improved, scores in wards 2, 7, and 8 have declined significantly. Several schools in wards seven and eight have seen startling declines in test scores of 15 to 20 percentage points or more. Ward 7, for example, declined 12 percentile points between 1991 and 1996. The low achievement levels attest to the fact that thousands of children, especially in wards 2, 7, and 8, are not being taught the fundamental skills necessary to succeed after they leave DCPS.

example, declined 12 percentile points between 1991 and 1996. The low achievement levels attest to the fact that thousands of children, especially in wards 2, 7, and 8, are not being taught the fundamental skills necessary to succeed after they leave DCPS (District of Columbia Financial Responsibility and Management Assistance Authority , *1996*, Failure of Public Education in the District section, para. 10)

Then, the report concludes that the poor performance derives from ―the school system's failure to fully prepare them for test taking and teach them critical thinking and analysis skills" (District of Columbia Financial Responsibility and Management Assistance Authority, 1996, Failure of Public Education in the District section, para 11).

The DC School Board's authority was stripped. A nine-member District of Columbia Public Schools Emergency Transitional Education Board of Trustees (Board of Trustees) was established until June 30, 2000 to oversee the district, and the superintendent reported to this Board. Soon, a new superintendent, Becton, was appointed. Becton appointed a new Chief Academic Officer, Ackerman, to work on improving academic performance. She later became the Superintendent when Becton resigned in 1998. Both Becton and Ackerman implemented measures to make schools accountable for student performance, and an assessment, the SAT9, became a high stakes test for principals, schools, and students.

The new administration issued policies that promoted reform-oriented instruction and policies that made schools accountable for students' academic performance. Becton created a strategic plan called ―Children First" to implement DC Goals 2000 (District of Columbia Public Schools, 1997). The instruction described in this document was reform-

oriented instruction.

At the same time, Becton and Ackerman introduced the following policies that made the district assessment high stakes. First, the district switched state assessments from CTBS to SAT9, and reading and math test scores on the SAT 9 during the spring 1997 were used as the baseline scores to measure improvement of each school (District of Columbia Public Schools, 1997). The district also planned to create a school profile that would show test scores, improvement, programs offered at each school, and community information, which was completed much later than originally planned because the district's record keeping was poor.

Second, the district shortened the duration of each principal's contract from three years to one year. Improvement of academic performance, which is mostly measured by the improvement on SAT 9, became a large part of principal evaluation (District of Columbia Public Schools, 1997). Third, the district introduced a strict promotion guideline for students. Students who performed at "below basic" on SAT 9 were to be held back (District of Columbia Public Schools, 1997). Fourth, the district proposed introducing performance pay for teachers (Mathews, 1998, Nov 18). Fifth, the district proposed reconstituting low performing schools[9]. Although the last two proposals were not completely implemented, this series of policies suggests that math and reading SAT 9 became high stakes for both students and schools.

Implementation of target assistance to low performing schools reinforced the

[9] In June 1999 the control board finally gave preliminary approval of the labor contract that linked some raises of teacher salary to performance. The criteria were to be decided between the superintendent and union leader, and it was planned that the performance of schools would be based on improved student achievement, learning climate and safely and other factors. (Wilgoren, Jun 10 1999 Washington Post)

notion that SAT 9 reflects whether schools need additional support. During the school year of 1996 -1997 schools that showed a decline on the state test for the past three years were identified, and the teachers at those schools received targeted assistance, which included training from the Lab for Student Success (LSS) at Temple University. Schools also received a full time ―change facilitator" to augment assistance provided by LSS. Later, these schools introduced Comprehensive School Reform programs[10]. These policies suggest that the test results are the measurement of student learning, that schools need to improve student performance on the test, and that the test is to be used for accountability.

Other actors in the field of DCPS who promoted the institutional logics of teaching for the test (although the original goal was not changing instruction to teacher-centered drilling but to make schools accountable for improving student learning) were parent groups. Parent United in DCPS often voiced concerns about school conditions and published school-by-school results of CTBS administered in the spring of 1996 with the amount of federal Title One funding awarded to each school in the spring 1997. The publication by Parent United stated:

> Since the CTBS itself was a poor measure of achievement because as a test it was out of date, and measured low level skills, students who did poorly were worse off than they thought. Thank goodness we are using SAT 9 now. The results however will not be available until sometime this fall, so planning of next year's

[10] The district did not select one program. Schools introduced different programs.

program must use CTBS results. (Parent United for DCPS, 1997, How Are We Doing: Looking at CTBS Scores Section, para4)[11]

This statement indicates that in 1997, Parent United, a parent group, viewed SAT 9 as a reliable measurement of the quality of education, thus the parent group promoted the idea that performance on SAT 9 reflects the improvement of student learning.

Review of the documents suggests that schools and teachers might have received concrete information about questions asked in SAT 9. The district tested students on SAT 9 twice, in the fall and spring, to measure the improvement. Starting the second year, schools were allowed to use test preparation materials. However, I could not identify documents that were suggested by the district to be used for the test preparation. The only sources I could identify were articles from the Washington Post, which reported schools were using workbooks and item batteries that appeared on the test. The Washington Post also reported that schools had stopped regular instruction to cover content that appeared on the test. The district adopted only the multiple-choice portion of SAT 9 (Welgoren, April 6, 1998). These descriptions suggest that schools and teachers saw that instruction promoted by the test was teacher-centered instruction that focused on mastery of skills that would be tested on SAT 9.

The district maintained the claim that academic standards and SAT 9 were aligned. As stated in one document created by the district, SAT 9 is to measure whether students meet with the district standards (District of Columbia, no date). However, an evaluation report produced by an external organization indicated that students were not

[11] Although SAT 9 was considered to be measuring higher order thinking, the district did not administer a writing portion of the test, and later the U.S. Department of Education concluded that SAT 9 does not measure higher order thinking.

meeting the district standards (Council of Great City Schools, 2004).

While a series of policies promoted the idea that student performance on the SAT 9 was important and that schools and teachers could see the problems that appear on the test, reform-oriented instruction was not carried out with this level of strength and specificity. The district promoted reform-oriented instruction through different carriers, i.e., strategic plans, standards, and curriculum. The strategic plans created during the late 1990s repeatedly stated that the district would align standards, curriculum, and assessment. A small number of teachers received training from organizations to learn math and science instruction that stressed hands-on, real life application and worked to create a new assessment in science in 1996. In 2002, the district completed their creation of academic standards that were in line with National Education Standards and the Improvement Council's ideas of performance standards. This is an organization created during the Clinton Administration to certify state content and student performance standards. The district's math standards prescribed activities and approaches promoted by the NCTM standards and encouraged integrating academic subjects (District of Columbia Public Schools, 2002). The strategic plans created by the district mention that students should learn critical thinking and that students should be assessed by performance assessment and other assessment.

Science curriculum introduced during this time period was based on reform-oriented science instruction. The district introduced FOSS for science, which is based on the idea that students learn science through hands-on activities (Foss Project, no date). Within these policy documents, the teacher's role was viewed as designing activities that facilitate problem-solving, activities that are carefully designed to be meaningful for their

students. Recommendation for integrating academic subjects also suggests that the academic standards imagined teachers to be knowledgeable about student learning.

> –By looking at the essential skills from different disciplines, teachers can see an overlapping of skills which makes integration of content feasible. Making these connections helps to create many opportunities to extend learning and to provide more enriching experiences for both teachers and students" (District of Columbia Public Schools, no date).

However, the academic standards were vague in their description of content and performance standards[12]. For example, the performance standard of –students solve non-routine problems" repeatedly appears throughout the math standards without giving specific examples of what types of non-routine problems they should be able to solve, and how these skills should progress. The standards did not clearly state what exactly students should learn at each grade level and how skills should develop over time. Although the district's strategic plan states that students should cultivate critical thinking (District of Columbia Public Schools, 2001), there is no document produced by the district to provide guidance as to how teachers could design activities to promote such critical thinking. The district's guide to standards and the pacing guide state that teachers should use multiple forms of assessments, including –open ended assessments, multiple choice assessments, or products, performances or processes" (District of Columbia Public Schools, no date). However, according to the articles that appeared in the Washington

[12] Performance standards indicate both the nature of the evidence required to demonstrate that content standards has been met and the quality of performance that will be deemed acceptable.

Post, the emphasis on the SAT 9 test was more widespread than using multiple assessments.

The review of the documents suggests that policy pressures that promoted reform-oriented instruction were less concrete than policy pressures that promoted teaching for SAT 9. In addition, the document I reviewed did not indicate that there was a regulative pressure on the schools to comply with the standards and curriculum. As it seems, the policy pressures to teach for the test were more persuasive than policy pressures that promoted reform-oriented instruction.

6.3.2 Governance

Site-based management continued during this time period. From 1996 to 2000 the district was under the oversight of the Congress, and DC's superintendents reported to the Control Board. Site-based management was not questioned, and the district leadership strongly supported continuing with it. Site-based management was viewed as the best way to improve student academic performance. For example, Superintendent Ackerman made it her priority to get parents' input on the school budget, and the Washington Post reported that Ackerman's effort for involving parents in this process is based on the following belief: "School communities know best what their children need and can manage their own schools without much interference from downtown" (Strauss, May, 30, 1999, The Washington Post).

Site-based management provided schools with autonomy with regard to instructional decisions; hence, schools were not required to follow the district's

instructional policy, curriculum, and pacing guide. Later, an evaluation report concluded that this site-based management contributed to the lack of district-wide instructional policy (Council of Great City Schools, 2004). At the same time, schools had to be accountable for students' improvement on SAT 9. The district's evaluation of schools and principals based on SAT 9 scores continued, and the Washington Post continued to publish the performance of schools' scores on SAT 9. The district required students who did not perform well on SAT 9 to attend Summer School. The schools were identified as in need of targeted assistance based on SAT 9 scores, and the teachers at those schools had to reapply for their positions.

6.3.3 Spread of Institutional Logics

It was not clear to what extent the logics of teaching for the test spread in the classroom, as there was not a systematic study of it. The documents I could find were newspaper articles and a report created by a non- profit organization in which teachers in neighborhood schools described their instruction. These documents suggest that schools taught for the test, although some teachers resisted this approach. The description in the newspaper suggested the instruction was teacher-centered, worksheet-based, and involved drilling for SAT 9. During this time period, The Washington Post repeatedly reported that teachers were teaching for the test. It reported that:

> –SAT 9 is helping to set the style and pace of instruction in DC schools. Having given their students a battery of SAT 9 reading and math tests in September and

October, teachers at all 146 schools are focusing intense effort on the SAT 9 barrage set for the spring." (Mathews & Strauss, 1998, Nov 23, 1998,).

Many schools ~~t~~tailored their lessons to materials covered on the test" (Welgoren, The Washington Post, April 6, 1998). The Post also reported that schools were dropping arts and music classes in 2000 (Bennings, 1999, July 5, 1999; Blum, 2000, June 23, 2000;).

Newspaper articles and the evaluation report of a particular neighborhood published during this time period suggest that although teachers might have taught for the test, they were skeptical *about* the test. First, the teachers did not see the alignment between SAT 9 and curriculum and standards. A community organization conducted a focus group of teachers in one neighborhood and found that the teachers were pressed to teach for the test that did not align with the district's standards (DC VOICE, 2001). Second, some teachers did not see the categorization of student performance as a reliable measure of student learning. The Washington Post conducted a survey of DC area teachers and follow-up interviews with selected teachers and reported that some teachers were in doubt of the district's strong emphasis on measuring performance improvement by the categories set by SAT 9 (i.e. below basic, basic, proficient, and advanced). The teachers reported that the district's emphasis on moving students from one category to another was too broad because students' performance within the same category varied, and this categorization did not reflect improvement of individual students if they stayed in the same category (Benning, Washington Post July 5, 1999). Third, some teachers did not see the assessment as the way to improve their instruction. The Post reported that some teachers commented that ~~l~~low test scores mostly from meager school resource and

students who are raised without respect for teachers or motivation for study, and it is not our fault" (Benning, Washington Post, July 5, 1999).

The newspaper reported that some schools that served students from affluent families made some effort to limit test preparation because parents protested test preparation. Parents from affluent areas boycotted schools[13] to protest the test-driven instruction. According to the Washington Post, these parents supported using standardized testing to help determine student progress, but they opposed classroom instruction becoming test-driven. The parents argued that ~~t~~eachers and principals still need training in how to reshape classroom instruction so that by doing regular teaching well during the school year, they will have prepared students for the test" (Welgoren, The Washington Post, April 6, 1998). The district maintained that most schools kept a balance between test preparation and regular teaching (Welgoren, April 6, 1998).

Taken together, the available documents suggest that pressure for teaching for the test were present in schools, and at least some parents and teachers were aware that classroom instruction was test-driven. However, I could not find any documents that inform how reform-oriented instruction spread to schools. The evaluation team of the Council of Great City Schools concluded that there had not been a district wide instructional policy and strategy (Council of Great City Schools, 2004).

[13] In 2001 May, The Washington Post reports ~~a~~ relative handful of parents—based in affluent NW Washington—attempt[ed] a similar boycott of last month's exams" because they see the strong emphasis on testing as a waste of students' time, and other subjects are cut.

6.4 FROM 2004-2008: MOVING FROM REFORM-ORIENTED INSTRUCTION TO TEACHING FOR THE TEST

Between 2004 and 2008, the organizational field of DCPS is characterized as the centralization of instructional policy. The district attempted to promote a uniform message about good instruction. Good instruction was viewed as reform-oriented instruction, and standards, curriculum, and assessment were to promote reform-oriented instruction. The difference from the previous time period was that the district abandoned site-based management. During this time period, the district required all schools to follow the district standards and curriculum.

In order to comply with the requirement from NCLB that the assessment and the standards must be aligned, the district created an assessment, DC CAS (DC Comprehensive Assessment System). The district also created new standards. From 2005-06, the district promoted reform-oriented instruction through professional development on the new standards. However, when the high stakes assessment required by NCLB was administered, professional development and materials associated with DC CAS promoted a different concept of instruction. This section presents how policy pressures associated with the standards and assessment resulted in promoting different institutional logics about good instruction and the use of assessment.

The outline of this section is as follows. First, I present the concept of good instruction promoted by the new standards and DC CAS to show how policy pressures associated with the standards and DC CAS resulted in promoting different conceptions of

good instruction. Second, I present how the change of governance shaped the concept of good instruction and the use of assessment promoted by DC CAS and the new standards.

6.4.1 Institutional Logics: From Reform Oriented Instruction to Teaching for the Test

Between 2005 and 2008, the district attempted to promote standards-based systemic reform by aligning standards and assessment. However, the review of the documents and observation of professional development in 2008 suggests that policy pressures associated with standards and DC CAS promoted different conceptions of good instruction and how assessment should be used.

Below, I describe how this change happened. First, I describe how the district promoted the institutional logics of teaching for the standards. Second, I describe how policy pressures promoted logics of teaching for the test. The table below presents policy initiatives during this time period. In this section, I draw upon my observation of district sponsored professional development sessions in addition to a review of documents.

Table 7 Timetable of Policy Initiatives between 2004 and 2008

Year	Policy Initiatives
2004	―Restoring Excellence to the District of Columbia Public Schools" by Council of Great City Schools. DC Compact created and community meetings and roundtables for standards begin.
2005	―Declaration of Education" by DCPS District began a series of professional development sessions for the standards. Math and Reading Standards were implemented in 2005-06 School Year New textbooks were purchased for 2005-06 School Year

2006	Science and Social Studies Standards were implemented in the 2005-06 school year. –Strategic Education Plan" by DCPS District had Saturday Schools for four Saturdays. DC CAS was administered first time in spring 2006 DC BAS (benchmark test) began in 2006-07 school year US Department of Education indicates the district as –high risk" for managing federal funds.
2007	Jan 2007 Mayoral control begins. Chancellor Rhee begins her term. New Science textbook was introduced in 2007-08 School Year
2008	The district started 4 months of Saturday Schools

Between 2004 and 2006, the district and other stakeholders promoted standards-based reform. The new standards and curriculum promoted reform-oriented instruction. The report by the Council of Great City Schools, *Restoring Excellence to the District of Columbia Public Schools*, offered a clearer direction of how the district should implement the standards-based reform. This report identified the problem of low academic performance among DCPS students as deriving from the poor quality of the district's academic standards and delivery, and recommended establishing a district-wide instructional policy that was in line with standards-based reform. The report concluded that site-based management had contributed to the poor delivery of district's instruction policy. The report states:

> –In short, the DC school district has abdicated its leadership responsibility for student achievement to the schools and has had trouble hitting its instructional mark over the years because so many people were aiming in different directions. The result is what one sees today: No plan for improving student performance,

low expectations for children, no accountability for results, haphazard instruction, incoherent programming, and dismal outcomes" (The Council of Great City Schools, 2004, p.10.).

Taking the recommendations of this report, the district created new standards, adopted a new curriculum, and developed a new assessment, DC CAS. The district created the new standards with strong public support. Various community organizations were involved in the discussion of the new standards. Although the philosophy of math and science instruction did not change from the previous time period, the district leadership portrayed the new standards as the break from the past standards. The new standards were clearer and more explicit about the content, skills and activities in the instruction than the previous standards. The Washington Post reported that new standards "make it easier for teachers at schools to implement…they will have a clear direction on what the district is aiming for" (Haynes, The Washington Post, August 25, 2005). The district claimed that all these efforts were to "dramatically change classrooms" by "making instruction more engaging, lessons more rigorous and textbooks more relevant" (Haynes, 2005, March 6,).

Below I present policy pressures that promoted reform-oriented instruction. The district completed math standards before science, and there were more activities associated with math standards than science standards.

Math

Reform-oriented instruction was promoted through professional development for standards and the new curriculum, *Everyday Mathematics*. Although both the standards and curriculum shared the same view of how students should learn mathematics, they promoted different concepts about how teachers should plan instruction.

Although the old standards and new standards both promoted the instructional approaches recommended by the National Council of Teachers of Mathematics (NCTM), the district officials portrayed the new standards and curriculum, *Everyday Math*, as new and believed that the new standards would change instruction dramatically. The new standards were more explicit about the content, depth, and activities the students should learn than the older standards. The standards began with the description of the underlying philosophy of mathematics teaching and learning that was adopted from Massachusetts Standards. The standards stated that students' learning should ~~–g~~o beyond acquiring procedural mathematical skills," and that problem solving is the process and outcome of math instruction. Students need ~~–m~~any opportunities to formulate questions, model problem situations in variety of ways," and ~~–g~~eneralize mathematical relationships," ~~–e~~xplore connections among concepts," and ~~–s~~olve problem in both mathematical and everyday contexts." The role of teachers was to encourage students to ~~–t~~alk, reflect and use multiple problems strategies" (District of Columbia Public Schools, 2005, pp. 3-4). Students were to use multiple representations of mathematical concepts and make convincing arguments.

Professional development by the district focused on helping teachers to implement standards-driven-instruction in their classrooms. The district's professional development used ~~–p~~ower standards" and ~~–u~~nwrapping standards" to help teachers

understand the standards and design instruction (Center for Performance Assessment, 2004, pp. 14-15). This approach was based on the idea that not all standards are equally important, and teachers could structure their classroom instruction around "power standards." Power standards were the fundamental skills and concepts to be mastered by the end of each grade level, and they included concepts and skills that students needed to progress to the next grade level. Other standards were considered as connected standards that go along with the power standards. The teachers could cover them as they teach power standards (Center for Performance Assessment, 2004). In short, within the limited instruction time, power standards were the must-be-learned standards.

This professional development that focused on power standards and connected standards purported the idea that teachers should plan lessons around power standards. During the mandatory professional development, teachers identified power standards by looking at a vertical alignment of skills and concepts across grade levels and across academic subjects. Teachers also identified the connecting standards that could be covered as they teach the power standards. Then, teachers reviewed the power standards to identify what skills and concepts should be taught to meet the power standards. The teachers looked over the power standards for previous grades to understand how students learned related skills in previous grades. In the session, teachers created lessons to teach power standards. The lesson plans included engaging scenarios, essential questions, and performance assessment. These activities reflect the idea that teachers should create activities around power standards, rather than merely following textbooks. The professional development also taught that classroom tasks should include various levels of skills, from low to high levels of thinking. The teachers identified different cognitive

tasks by using Bloom's taxonomy (Center for Performance Assessment, 2004). These activities suggest that the teacher's role was to identify when and how power standards should be taught and determine the best activities for students based on students' academic standing and differing learning styles. The activities teachers learned during the professional development suggested that good math instruction includes math problems that facilitate multiple tasks from low to high levels of thinking and have an interdisciplinary nature. The role of the teacher is to decide when the power standards should be taught to their students based on knowledge about how and what content and skills students have learned previously.

Although *Everyday Mathematics*, the curriculum adopted by the district, was also based on the idea that students should learn math concepts through math problems that require multiple math tasks, the curriculum had a pre-set sequence and math problems. Teachers attended professional development sessions for *Everyday Mathematics* and learned that *Everyday Mathematics* enacts a spiral structure, and they were encouraged to follow the prescribed curriculum as much as they could.

Both *Everyday Mathematics* and the standards were based on the idea that students should learn math through working on a problems related to their everyday experiences. The Washington Post reported the characteristics of new curriculum as follows.

> New math curricula and textbooks are designed to help students from kindergarten to 12th grade master their subject better by learning how key concepts apply to life situations rather than by using the age old approach of

relying largely on rote memorization, drills and worksheets[14]" (Haynes, 2005, August 25,).

The district official explained that the new mathematics curriculum was ―aimed at making math more relevant, in the past we used rote memorization to teach math, while that is still valuable, what is more important is that students will understand how the math concepts connect, how they apply to life situations. With the new approach…these concepts are going to stay with the students" (Haynes, 2005, August 25, p. T 10).

The district maintained that its curriculum and standards were aligned, and students' academic performance would improve by aligning standards, curriculum, and professional development[15]. However, the district's math standards and curriculum, *Everyday Mathematics*, carried different messages about how a teacher should plan lessons. Although *Everyday Mathematics* and the district standards both promoted instruction recommended by NCTM standards that promoted a conceptual understanding of math by making students work on real world math problems, *Everyday Mathematics* was a highly prescribed curriculum. On the other hand, the professional development focused on the standards promoted teacher-designed instructional activities around power standards.

---

[14] During the time the district was selecting a math curriculum, some math experts sent an open letter to request the district to drop *Everyday Mathematics*. The letter stated, ―the real problem is that the core skills and key topics are not taught. They are touched upon but not given nearly the amount of time needed for mastery" said Bastian Braams, visiting associate professor of math and computer science at Emory University, who has been pressing the DC School Boards to drop the textbook. The district responded to this criticism by saying, ―kids don't come away loving math with the traditional approach. We lose them. They get discouraged from taking advanced courses… the research on the materials we adopted is excellent. Within two years, kids are turning around their achievement and linking math enough to go on to advanced courses" (Haynes, The Washington Post, August 2005 p.T 10).

[15] ―standards are basis for curricula, previously classroom instruction often was unrelated to the standards and to the achievement test given to students. The only way we are going to get better test results is to make sure the textbooks, curricula and the PD for teachers are all linked to the standards."

One evaluation study conducted by a non-profit organization during the school year 2005-06 reports the consequence of this gap. This non-profit organization conducted the first year implementation of the standards and curriculum through a focus group with selected teachers in one neighborhood, i.e. a high school, a feeder junior high, and elementary schools. In the focus group, some teachers reported that their principal had told them only to use *Everyday Mathematics* but not other materials associated with standards; teachers in other schools report that they created their curriculum based on the standards (Community Organization, June 2006)[16]. As it is, teachers differently weighed the importance of standards and curriculum.

During the school year 2005-2006, the district continued with professional development to encourage teachers to teach for the standards. The district seemed to be aware that the professional development and the guidance from *Everyday Mathematics* conveyed disparate ideas about how teachers should teach. In a professional development session in January 2006, the district sent out a memorandum that stated that performance tasks for performance assessment were not required for all lessons, and the teachers should use performance tasks only when they teach power standards (District of Columbia Public Schools, 2006a). During this time, the district also sent out forms to pressure schools to teach for the standards, such as an affirmation form in which teachers were asked to sign their name under the statement that they would teach for the standards, a form including guidelines about the time that should be spent on reading and math instruction, and outlines of how a standards-based classroom should look (District of

[16] Also more than half of the teachers who participated in the focus group reported that they did not receive curriculum materials before school started (Community Organization, June, 2006).

Columbia Public Schools 2005a, 2005b, 2006). The district also created a pamphlet for parents. It explained what their children would be learning according to the new standards and how parents could help support their children's learning (District of Columbia Public Schools, no date). In summary, the standards-based instruction was promoted through professional development and guidance that described what a good classroom environment should look like, and an affirmation form that required teachers to sign their names, making a commitment to change classroom instruction to standards-based reform-oriented instruction.

Science

The concept of good instruction and the role of teachers had not changed in the new science standards. Various stakeholders, including teacher representatives, district officials, representatives from community organizations, and parents, participated in roundtable discussions to build consensus on the overall direction of the standards. The group decided to draw upon Indiana's science standards. The new science standards were implemented in the school year 2006-07. The only difference between the new and previous standards was that the new standards clarified skills and concepts to be learned at each grade level. As with the math standards, the science standards had an introductory section, "guiding principles of science education," which was partly adopted from the Massachusetts Science and Technology Engineering Framework and explained the goals of science education, including how science teaching and learning should look. It stated that science knowledge is constantly changing, and students should connect

multiple concepts as they learn science. It also stated the importance of science literacy and students making practical applications through technology. The standard states that science education should nurture and sustain students‘ curiosity about their environment so that they can acquire a solid scientific foundation. Science instruction was intended as an active process of investigation, a critical review of evidence, and an encouragement of application of knowledge beyond the bounds of science to support learning goals in all subjects.

Science teaching and learning were viewed as ―investigation, experimentation, and problem solving" and ―central to effective science education." ―Students need opportunities to talk about their work in focused discussions with peers and with those who have more experience and expertise" (District of Columbia Public Schools, 2006, p. 4). This view of science teaching is in line with existing standards and the existing curriculum, FOSS kit, which facilitates students‘ learning about scientific laws through hands-on activities, experiment, and observation (FOSS Project, No Date); however, the content of the FOSS kit did not correspond with the content area of the new standards. The district personnel reported that the teachers had to supplement with other materials to cover the content of the standards.

There was less intensity of institutional pressure to spread the concept of instruction in science than in mathematics. The professional development for the science standards was only one day. The district did not introduce a new textbook with the new standards, and the district did not provide guidance on what materials teachers should use. Rather, the district allowed teachers to use any materials they wished in order to meet the standards. While the School Board approved the standards to be implemented in

2006-2007, the district did not order a new textbook for science and the Washington Post reported the following reasons provided by the district. :

> ~~t~~eachers will use current textbook and incorporate new material into their lesson plans and add field trips to history and science museums. There are not huge gaps between the new standards and old textbooks…it is not like the textbooks will be the sole deliverers of quality instruction. The power of the textbooks comes out in the quality of teaching." (Haynes, & Labbe, 2006, August 24).

The district purchased a new textbook in the school year 2007-08. The new textbook was presented as one of many resources teachers could draw upon.

In the beginning of the school year 2006-07, the district provided professional development that followed the approach taken for math, including teaching about power standards and connected standards. In the professional development, the teachers learned the connection between standards, power standards, and the assessment, DC CAS. The policy message the district intended to convey was that standards and assessment were aligned because power standards would be tested on the assessment. The session lasted approximately a half day, much shorter than the sessions for the math and reading standards.

DC CAS for science was field-tested in the spring of 2007 and was fully implemented in the spring of 2008. The year I collected data was the first year of full implementation for the science test. It was administered only to 5th graders. In January 2008, the district distributed a ~~b~~lue print," a booklet that listed standards to be tested and the number of points given to each question that was attached to each standard. The information about testing was distributed through the science chairs' meetings that

happened every couple of months. However, not all schools participated in the chair‘s meeting. The benchmark test was not administered for science. Activities associated with testing were far fewer for science.

Although the assessment was supposed to promote one unified concept about good math instruction, policy pressures associated with DC CAS resulted in promoting different concepts of how teachers should teach. The policy pressures associated with DC CAS promoted the idea that teachers should teach for tested standards, rather than power standards. This conceptualization became clear during the school year 2007-08. While policy pressures associated with DC CAS between 2005-06 rarely touched upon how teachers should prepare their students for DC CAS, the district began to provide more guidance during 2007-08. The district provided professional development, benchmark tests, and a blue print for DC CAS. These policy pressures highlighted the importance of improving performance on DC CAS. These policy pressures also presented information on what problems would appear on the test. The conception of instruction promoted by policy pressures associated with DC CAS was that teachers should teach for tested standards rather than teach for power standards.

Although the district originally promoted the logic that students‘ academic performance would improve by aligning standards, curriculum, and assessment, professional development associated with DC CAS began to promote different conceptualizations of good instruction and different ideas about how teachers should plan

lessons. The policy pressures associated with DC CAS promoted the idea that teachers should teach for ―frequently tested standards," and the district used benchmark tests, DC BAS, to help teachers identify the standards. This conceptualization of good instruction was different from the concept of teaching for power standards. The role of the teacher for planning instruction changed from identifying power standards and relationships among supporting standards and creating lesson plans around them with complex tasks and performance assessment to teaching frequently tested standards. The district did not send policy messages about how to teach for frequently tested math standards. The carriers of the teach-for-the-test concept were more concrete because they presented *what* to teach. The carriers also included materials teachers could use in the classroom. There were more frequent regulative pressures to promote this logic rather than the logic of teaching for standards. In general, while the logic of standards-based teaching was promoted through normative pressures (such as professional development that encouraged teachers to teach for the standards), the concept of teaching for the test was promoted through policy pressures that were frequent and that presented concrete information about what to teach. Below I describe how logics of reform-oriented instruction shifted towards teaching for frequently tested standards.

During early part of 2005-06, the activities teachers experienced during their professional development put forth the notion that assessment would be in line with the power standards. In the professional development sessions, in order to learn about standards in 2005 at each school, teachers did activities to identify power standards from the new standards, and each school submitted its own power standards to the district's

central office. Consequently, each school came up with different power standards[17]. In the professional development session that introduced the standards, teachers learned that power standards would appear on DC CAS because they are must-be-taught standards. However, the standards that would be tested on DC CAS as shown in the blue print, which was created by the test provider, were not necessarily the same as the standards teachers identified as power standards[18].

The review of professional development materials suggests that during 2005-06 and 2006-07, the district continued with professional development around standards, with gradual introduction of topics associated with DC CAS and tests. For example, during the school year 2006-07, professional development covered data use; however, the materials used in professional development did not provide information on how schools should use DC CAS data to change instruction. During 2006-07, the district assembled a group of math teachers to create standards-based worksheets that included power standards, essential questions, performance assessment, and state assessments problems from other states. These worksheets drew elements from the professional development offered the previous year, and they covered all power standards for math. This time, the district provided the teachers with power standards. However, this document did not reach all teachers. At some schools, teachers received the worksheets, but at other

[17] The materials from the organization that provided the professional development for the standards present the examples from other school districts. Other school districts created district-wide power standards. (Center for Performance Assessment, Making Standards Work, 2004)

[18] This non-alignment between standards and the standards that are tested derives from a lack of communication between the district's Department of Instruction, which was responsible for professional development, and the Department of Accountability, which was responsible for the state assessment. Although teachers were told that their power standards would be communicated to the district office, there was not much communication between these two departments, and the state test was created separately from the professional development (Personal communication with a non-profit organization) .

schools teachers did not know about them. Additionally, the district did not plan any professional development on how to use the worksheets. During 2006-07, the district began to administer interim assessment to predict students' performance on DC CAS; however, the district did not provide guidance on how these results were to be used.

During 2007-08, under Chancellor Rhee, the district's professional development began to focus on DC CAS. These professional development sessions promoted the conceptualization of good instruction as teaching for tested standards. These messages were sometimes explicitly stated in professional development. A series of policy pressures also promoted this conceptualization. For example, in the professional development on constructed responses in reading that I observed in January of 2008, professional development facilitators suggested to teachers that they should teach for the standards that are tested. At this professional development session, seven to eight teachers sat around a table with one facilitator for several hours. The teachers and the facilitator went over how to grade a constructed response. During this discussion, the facilitator repeatedly talked to the teachers about planning their lessons starting from tested standards. Below is an excerpt from my fieldnotes, which illustrate the justification the facilitator gave to teachers for planning instruction starting from the tested standards:

> —There are four curricula in your school: taught curriculum, learned curriculum, written curriculum, assessment curriculum. Taught curriculum is what you teach… learned curriculum is… Written is what is written (Note: in the) curriculum, and assessed curriculum is DC CAS. DC CAS is already telling you what are (on the) test. Let's now go back to the curriculum and you are teaching

(Note: from the test). In the standards environment, you start here (and the facilitator holds up the constructed response hand out)… I use the assessment to keep those other three curricula. As I move to teaching process, these curriculum, when I get the end I have this (meaning the standards that are tested)" (Fieldnote, January 2, 2008)

In the professional development session, teachers read the blue print and counted how many standards would be tested on the DC CAS. The district's central office informed teachers that instruction should focus on the tested standards, and that the district was not using power standards anymore. These activities associated with DC CAS minimized the concepts of instruction emphasized by the standards.

The professional development on standards encouraged teachers to identify power standards, to design engaging tasks, and to enable multiple levels of cognitive activity. These activities assumed teachers would design instruction by combining skills and topics across academic subjects and by taking into consideration how their students learn skills and concepts. Messages carried through the activities associated with testing did not portray this as the teacher's role. In the professional development, teachers were shown frequently tested standards, and they were told to teach for ~~f~~frequently tested standards." The focus of the professional development and the activities teachers experienced in the professional development directly contradicted the assumptions underlying the power standards.

The concept of teaching for the test was also promoted through policy pressures that carried explicit and concrete information about what teachers should teach. DC BAS, a benchmark test that was to predict how each student would perform on DC CAS,

began during the academic year 2006-2007. There were three DC BAS tests in a year. During the 2007-2008 school year, the district instructed principals on how to use the benchmark test results, and the principals provided the training to teachers at each school. DC BAS results indicated how many students were at each level, i.e., below basic, basic, proficient, and advanced. Teachers in the two schools I studied told me that the results came much more quickly during the year 2007-2008 than in previous years, and the information about student performance on the test was more detailed in 2007-08. DC BAS results showed which standards individual students missed and offered practice problems teachers could provide to students.

The importance of making AYP was communicated through other measures, and these measures became more consequential to schools as time passed. In addition to continuous publication of test scores of DC Public Schools in the Washington Post, the district's strategic plan published in 2006 categorized schools into incentive schools, target assistance schools, and restructuring schools, depending on their AYP status. Schools that performed well above AYP requirements were awarded an incentive award to provide programs of their choice, and they received more autonomy. During the school year 2007-2008, three schools that most improved student performance on DC CAS during the previous school year were awarded a bonus. Each staff member in these three schools received a bonus ranging from $2,000 to $10,000 according to their job title. The district's Saturday School was extended from several Saturdays in 2006-2007 to four months in 2007-2008. The district sent an invitation letter to attend Saturday School to students in schools that were in improvement status and who were close to scoring proficient on the DC CAS. The district also instructed schools to cancel field trips during

March so students could be prepared for DC CAS. All these activities carry the idea that schools should make AYP.

Although the district continued to use the math and science standards, policy pressures associated with DC CAS promoted the concept of good instruction as teaching for frequently tested standards. Categorizing schools by AYP status and rewarding school staff that made the most improvement in their DC CAS score promoted the idea that the DC CAS would hold schools accountable. In addition, DC BAS was to be used for improving instruction. A series of policy pressures associated with DC CAS, such as DC BAS and associated training, suggest that improved instruction means covering the standards that were missed on the benchmark test so that students would perform well on the next test.

6.4.2 Governance

The governance employed during these times was the centralization of instructional decision-making. Although the governing body of the education system moved from Board of Education oversight to mayoral control in 2007, the centralization of instructional decision-making continued.

The report by the Great City Schools in 2004 concluded that site-based management did not work to improve academic performance in the district. The report discussed that "standardizing the curriculum and professional development" is important to "improve the bottom line, student achievement" (The Council of Great City Schools, 2004, p. 16). The district centralized the decision-making about curriculum, standards,

and textbooks so that all schools would follow the district standards. Taking up the recommendations, the district began district-wide mandatory professional development sessions to introduce new standards in the spring of 2004 and continued providing professional development during the school year 2005-06. This is a significant change because *all* schools were required to participate in the professional development to roll out the standards on the same day, meaning most teachers came across the policy pressures that encouraged them to teach for the standards.

Although the district provided materials to encourage the implementation of the standards, the district did not monitor if these materials were used in schools. Furthermore, the centralization of instructional policy was mostly for math and reading. Schools continued with their own staffing models, which resulted in schools spending different amounts of time on science. The district instructed schools to teach reading and math for 90 minutes each, but it did not provide guidance for other academic subjects. Some schools had specialist teachers who taught science, while in other schools classroom teachers taught science. Some schools had art, music, and PE teachers, while other schools had only one of them[19]. In some schools, teachers integrated other academic subjects into reading or mathematics, while other schools had a set time for science lessons. Some schools had reading and math specialist teachers, while other schools did not have them.

[19] The central office announced a comprehensive staffing model to be implemented in schools in the 2008-2009 school year, so that all elementary schools would have art and music teachers as well as on-site coaches for reading and/or math.

Centralization and focus on academic performance was emphasized after the mayor took over the education system in 2007, and Chancellor Michelle Rhee began her term in July 2007. The new administration did not renew the contracts of principals who did not improve academic performance, or who did not seem to have plans to improve performance, which sent the message that improving test scores was the priority. Schools that were making AYP were called incentive schools, and they were allowed to introduce their own initiatives, such as math and science focus or art integration, while schools that were not making AYP received support from the district. Thus, schools that were performing well had more autonomy from the district's instructional policy.

It is clear then that educational governance during this time period was based on the idea that academic improvement should happen through strong district leadership on academic issues. However, the schools' staffing during 2007-2008 was not uniform, which made it difficult for some schools to receive instructional messages as they did not have teachers to attend district level meetings.

6.4.3 Spread of Institutional Logics

It is difficult to determine the degree to which the concept of good instruction and use of testing during this time period spread, as I did not find a systematic study about instructional practices across schools. From my interviews with non-profit organizations that were involved in the implementation of the standards, it seemed to me that a large number of teachers came across the term, "power standards," as the teachers were

required to attend professional development to learn the new standards. As I described in the previous section, some teachers received conflicting messages of whether they should plan lessons according to the standards or whether they should follow the sequence of *Everyday Mathematics*. The material to support implementation of the standards-based instruction was delivered to only a limited number of schools.

The concept of teaching for the test seemed to spread, particularly after the mayoral takeover. After dismal results on DC CAS during the first year, spring 2005, the district began benchmark assessment in 2006-07. However, not all schools received results in a timely manner, and the district did not provide professional development to all teachers about how benchmark assessments should be used. Stronger pressure was put on schools to use DC BAS results and other materials associated with DC CAS after the mayoral takeover. Professional development on constructed response reading was mandatory for all schools, except for the incentive schools in January 2008. DC BAS results were returned on time for the school year 2007-2008, and principals instructed teachers about how to analyze the test data at each school in January 2008. The district's central office trained principals about how to calculate their school's AYP requirement, so that the principals would know how many more students needed to be moved up to proficiency level to make AYP. An invitation from Saturday Schools for students who were close to passing the test was distributed to the schools, and teachers came to know about the incentive award through media and in teachers' meetings. The performance of DC CAS improved at the end of school year 2007-2008, and it was widely reported.

As for science, I did not find any documents that indicated the spread of the concept of teaching for the test among teachers and schools in science. When I collected

data in 2007-08, schools received the blue print, but I did not find any other policy pressures like the ones in math.

6.5 SUMMARY

In this section, I described the change of DCPS's organizational field by examining the change of institutional logics and governance. The review of the documents suggests that, between 1996 and 2004, although policies carried through district standards promoted reform-oriented instruction, policy pressures associated with high stakes testing promoted the idea of teaching for the test, which seemed closer to traditional teacher-centered instruction. The district promoted conflicting concepts about the role of teachers in planning lessons through policy pressures associated with standards and assessment after 2005.

The governance shaped the way policy pressures were carried to schools and teachers. Available evaluation reports and newspaper articles suggest that between 1996 and 2004 the site-based management prevented reform-oriented instruction from reaching schools and teachers. During this time, policy pressures associated with high stakes testing seemed to have reached schools and teachers. Also during this time, the use of assessment aimed to make schools accountable rather than improve instruction because SAT9 was not aligned with curriculum. The new standards and curriculum were more widely spread as the district provided mandatory professional development for all

teachers. Despite the district's attempt to promote a uniform conception of good instruction by aligning the standards, curriculum, and assessment in 2005, it began to promote different conceptions of how teachers should plan lessons. This change was accelerated in 2007 when the district shifted the focus of professional development to assessment. The teachers were instructed that they should teach for tested standards, rather than the power standards.

7.0 FINDINGS: POLICY PRESSURES ASSOCIATED WITH MATH AND SCIENCE IN THE FIELD OF DCPS

7.1 INTRODUCTION

Previous sections describe the change of the organizational field of DC‘s public schools from 1989 to 2008. Improving academic performance has been the underlying logic around which various instructional policies were created. Although the idea was to align the standards, curriculum, and assessment, policy pressures associated with the standards and assessment resulted in promoting different conceptualizations of good instruction.

In this section, I first describe how policy pressures associated with DC CAS differ between science and math because one of my inquiries was about how high stakes testing and low stakes testing changed teachers‘ classroom instruction. I describe how the nature of policy pressures associated with DC CAS differed between math and science. Second, I describe the district‘s instructional policy context during 2007-08 because I found existing instructional practices at the school level also shaped how teachers responded to policy pressures associated with math and science assessment.

7.2 DIFFERENCE IN POLICY PRESSURES

Various differences in math and science testing shaped the way teachers came across policy pressures as well as how they incorporated them into their instruction. Math and science testing differed in the following three ways: a) stakes attached to the test, b) intensity of policy pressures, and c) the specificity of messages policy pressures carried. The table below lists policy pressures associated with math and science DC CAS. Below, I describe how these policy pressures differed in the intensity and specifications of the policy messages.

Table 8 Policy Pressures Associated with DC CAS

Year	Math	Science
SY 2005-06	**First Year of DC CAS Math** • Blue Print (list of standards to be tested, number of points for each standard and sample questions) • District provided sample questions (Test from other states)	
SY 2006-07	• DC BAS (benchmark test) • PD on Math Constructed Response (voluntary) • Blue Print • Saturday School	**DC CAS Science Pilot Test** • District provided sample questions (Test from other states)
SY 2007-08	• DC BAS (benchmark test) for Math • PD on Math Constructed Response (voluntary) • Blue Print and District-wide PD on Blue Print (Math and Reading) • PD on DC BAS analysis at each school • Saturday School (From January to April) (Math and Reading)	**First Year of DC CAS Science** • Blue Print (list of standards to be tested, number of points for each standard and sample questions)

Although both math and science tests were mandatory, the pressures regarding student performance was quite different. While math test results were counted in schools‘ AYP statuses, which consequently influenced each principal‘s contract renewal and each school‘s improvement status, science was not counted to AYP. The difference in the stakes attached seemed to contribute to the difference in policy pressures. The policy pressures associated with math assessment included the monitoring of student performance, but there was no such policy pressure associated with science.

Second, there were few policy pressures that carried the message that the teachers should teach for the test in science. The teachers received DC CAS results for math, but the teachers did not recall receiving the results for the science pilot test. The teachers periodically saw the results of math DC BAS, Benchmark Assessment System; however, the district did not administer the benchmark test for science. The teachers came across policy pressures associated with math testing more often than with science.

Third, the policy pressures associated with math carried more specific and detailed messages than the ones associated with science. The Teacher Report of benchmark assessment, DC BAS, which the teachers received three times during the school year 2007-08, showed the performance of individual students according to content area of math and reading. The Teacher Report also connects specific problems on the test to content areas. Hence, teachers could see which problems individual students missed, and which problems corresponded to which content areas and standards. There was no similar information available for science.

In summary, the difference between high stakes math and low stakes science tests has to do with the intensity of policy pressures the district sent out to schools and the

degree of concreteness of the information carried by the policy pressures. In the next section, I present the district's instructional policy about math and science because these policies also shaped school level math and science policy at the two schools I studied.

7.3 DIFFERENCE IN MATH AND SCIENCE INSTRUCTIONAL POLICY

In the DC public schools, the guidance from the district about instructional practices and the availability of resources to support teachers' plans for instruction differed between math and science. Although mathematics instruction was mandatory and schools were required to allocate 90 minutes every day for mathematics, there were no such guidelines for science. It was totally up to schools how science instruction would be structured. Although the teachers received the new textbook at the beginning of the school year 2007-08, the district did not have a set curriculum for science. The guidance from the district was that teachers should teach for the science standards, and teachers should design instruction by drawing resources that best fit their students and standards, including FOSS kits and the new textbook. The following account by the Science Director summarizes the district's approach to how teachers should plan science instruction.

> –Our curriculum is our standards… It is really about standards, and [teachers] really have to look and see what it is they need to teach, what concepts and skills they need to teach and they use the materials, that best teach those particular concepts and skills. We should be able to use whatever materials we have to

teach the standards instead of sticking to textbook and teaching straight from the textbook or even just straight from the kit" (interview with Science Director, on April 15, 2008)[20].

In contrast, the district adopted *Everyday Mathematics* as the district curriculum for math, although the implementation of it varied by school.

There were also differences between math and science in terms of the availability of materials to support implementation of the standards and curriculum. The district created a Standard-Based Worksheet and a pacing chart for math to help teachers implement the math standards, but the Standards-Based Worksheet for science was not completed. Although math standards present skills and concepts in the order of learning, science standards present topics, concepts, and skills in no particular or set order.

7.4 SUMMARY

This chapter presented the difference between math and science testing and the difference in the district's guidance for math and science instruction. Schools and teachers received more intense policy pressures that included concrete messages about instruction related to math testing than they received about science. The instructional contexts of math and science differed. While the district provided more concrete guidance about the

[20] This approach is also supported by the district's views of how students should learn science. "The idea behind kit was that we really want to focus on interactive hands-on aspects of science, which is always the main way students learn the skills and concepts in science." "Science should be interactive…it should be interdisciplinary, that should be hands-on, that it should include scientific literacy…reading a lot of non-fiction" (Interview with Science Director, April 15, 2008)

curriculum and time allocated for math instruction, schools had more latitude about how to structure science instruction. In Chapter 9 and 10, I present how these differences in policy pressures associated with math and science testing, and differences in existing contexts of math and science instruction, shaped teachers‘ instructional change. The next chapter, Chapter 8, describes school context.

8.0 FINDINGS: SCHOOL CONTEXT

8.1 INTRODUCTION: TWO ELEMENTARY SCHOOLS

As institutional theory and Cognitive Approach to policy implementation suggests, my study finds that the organizational context shapes the way teachers respond to policy pressures. In this study, I collected data from a total of eight teachers at two schools: Jefferson Elementary School and Central Elementary School. The differences between the schools were not limited to student population and historical academic standing, but also organizational structure, school mission, staff assignment, how math and science had been taught, and the school's exposure to the institutional pressures originating from the district. Below I describe the differences between the two schools as a backdrop for the next chapter, which considers the study's findings of how individual teachers responded to policy pressures.

8.2 JEFFERSON ELEMENTARY SCHOOL

Jefferson Elementary School had a reputation of being one of the highest performing schools in the district for decades, and it had been making AYP since NCLB was

implemented. The school was praised for improving the academic performance of minority students. The school is in a neighborhood of middle class families and less than 3% of students were receiving free or reduced-priced lunch. 70% of the students were white, 20% were Black and the remaining students were either Asian or Hispanic. Teachers often mentioned that their students were high performers and that having good family support contributes to high performance. Parent involvement was high. The school‘s Parents Association organized many fund-raising events to support various student activities, such as a brass band concert, musical performances, and school fairs. There were enough parent volunteers for field trips and other activities.

The school promoted a student-centered approach through integrating arts and academic subjects in classroom instruction and school events. School events, professional development, and the principal‘s message to teachers suggested this educational philosophy. The school had many events where student work was showcased, such as a science fair, a spelling bee contest, musical performances, and other fairs where students presented their research projects and art work. For example, the fourth graders had a Suitcase Fair. Individual students selected one U.S. state and researched the state‘s major industry, geography, culture, and history. The students used graphs, pictures and essays to describe their findings, and displayed it in a suitcase. The students worked over two months on this project that integrated math, social studies, arts, and language arts components. The students were given a schedule of when each assignment should be completed, and the teachers provided feedback throughout the project, such as providing feedback on the first draft of the students‘ descriptions of their state‘s major industry. On the day of the Fair, the students brought one traditional dinner

of their state, and the suitcase was filled-up with various types of information about the state. Parents, students, and teachers visited the exhibition. The students served the traditional foods and explained their suitcases to the visitors. The goals, schedule, and grading rubric for each assignment and an example of a suitcase were on display when students worked on the project so that students could see what was expected from the project.

Jefferson Elementary School was one of a small number of schools in the district that had full-time art, music, and PE teachers. Students took each of these "special subjects" once per week. During this time, teachers had 45 minutes of planning time, 3 times each week. There was a morning recess from 11:00 AM to 11:20 AM when students played in the schoolyard or in the gym, and teachers took turns watching them. Students also had a 30 minute recess after lunch to play outside.

The school was an old brick building with an open space floor plan. It was located in a quiet residential area with tree-lined streets. The school's enrollment was increasing, and the school added several trailers as classrooms. When visitors enter the school building, there is a light filled open-space with high ceilings and many windows. On the bulletin board at the entrance, there were letters from U.S. presidents thanking the students for calendars that the Parent Association had created every year for fund raising, as well as thank you letters to students participating in national reading events. Banners hung from the walls, including the banner of blue ribbon school, art integration education program from performance theaters, and the school council's welcome banners written in English and Spanish. On the side wall at the entrance was a bulletin board, "We are Art School," and pictures of staff members.

Student art was posted throughout the building, and the art was replaced with new projects every one to two months. The art and music works exhibited often incorporated other subject areas‘ concepts. For example, the students created pictures with the theme of the Great Depression by re-creating art work produced during the Depression era by combining different methods, such as photography, oil painting, and pastel. Students added descriptions of how they created the work, and why they selected a particular object to represent the Depression era. Another time, there were portraits of friends drawn by the second graders. The principal explained to me that the second graders used a ruler to measure each part of their friends‘ faces to draw the pictures. Fifth graders composed music by playing instruments used by American Indians. The principal and teachers mentioned that art and music teachers often asked about what students were learning in other subject areas, and they incorporated these elements into their lessons.

Upstairs was a classroom space. As this school was built with an open space floor plan, there were partitions that divided open spaces into classroom areas. These partitions function as bulletin boards to display student work, such as graphs students created from data they collected, short sentences composed by students, or diagrams students draw as they learn about shapes. Some of the student work was accompanied by the district standards, such as ―5.DASP.2 Construct, draw conclusions, and make predictions from various representations of data sets, including tables, line graphs, line plots, circle graphs and bar graphs (where symbols or scales represent multiple units),” but not for all the work. The principal did not pressure the teachers to put standards on the wall.

Besides having resources to provide art, music and PE, Jefferson Elementary School also had subject specialist teachers who guided other teachers by translating policy messages from the district associated with standards and DC CAS. The school had two reading specialists and one math specialist. They worked with classroom teachers to provide support for everyday lessons, went over student performance data, and consulted with classroom teachers about what types of support they should provide for the students. The Specialists coordinated and provided individualized support to students who were lagging behind and arranged enrichment activities for high performers. They also provided small group activities to students who had difficulty working in a whole class setting or who might need more attention to catch up. The Specialists planned and provided professional development sessions for the teachers. For example, the two Reading Specialists provided voluntary professional development monthly sessions in the morning before school started throughout the school year, and most teachers attended to share and learn reading strategies. The Math Specialist facilitated sessions on error analysis of the DC CAS twice a year. One session lasted two hours and another session lasted one hour. When I was conducting interviews during the school year 2007-08, the Math Specialist mentioned that his role was changing. While he visited classrooms when they began implementing *Everyday Mathematics* in 2005-06, more of his time was spent recently on providing individualized or small lessons for students to make sure students in all groups can make progress in math. The Math Specialist described the reasons for this change as follows.

> —Because [testing has] gotten be so important, we (note: have to) show improvement with African American students, with English learning students,

non-native speakers…all these different categories, that is why [the Principal] is asking me to work with some of the groups or some of the individuals that are in danger of being left behind." (Interview with Math Specialist, Jan 8, 2008).

Having an in-house professional development capacity derived from the district's long history of site-based management allowed principals to make staffing decisions. In addition, Jefferson Elementary School was designated as a "high performing school" because its students had consistently excelled on the SAT 9 and DC CAS. The school received an incentive grant to be used for school initiatives and professional development. The principal decided to introduce an art integration program. With this award, Jefferson Elementary School became a part of the Art Integration Collaborative in the school year 2005-06 and began to receive professional development to provide art integration.

Rather than attending the district's professional development at the beginning of the school year and during winter break, teachers at Jefferson Elementary School participate in professional development provided by the Art Integration Collaborative. Besides the intensive training on the strategies for art integration, the teachers received on-going training from artists/instructors. The artists/instructors scheduled classroom visits with each teacher and they modeled art integration activities 8 times per semester, a total of 16 lessons in a year for each teacher. For example, one teacher had a musician in her classroom. The musician taught the concept of fractions by using music notes and drumming. The artists left materials, including lesson plans, so that the teachers could do the same activities next year. Jefferson Elementary School sent only a couple of teachers

or Reading and Math Specialists to district professional development, and they disseminated summaries of the professional development at school.

At Jefferson Elementary School, instructional practices and school activities were to promote instruction that was student-centered and interdisciplinary. This message about instruction is reflected in the activities, which are interdisciplinary in nature, such as learning about math concepts through art and music lessons, and the encouragement from the principal to promote art integration. Although the teachers said they felt the pressure of testing, they often added that the testing pressure is weak at Jefferson Elementary School. The principal told me that she did not want to say much about testing; rather, she thinks it is important to continue with the art integration program because she thinks it helps students learn better. However, it does not mean they ignore the state test. For example, the principal constantly informed teachers about the testing schedule and about what administrative work was associated with the benchmark test and state test. The teachers then created a committee to divide the administrative tasks associated with DC CAS. The principal sometimes mentions Jefferson Elementary School‘s performance on the district test and the assessment system the school had been using. The principal and grade level teachers met at the beginning of each school year to come up with goals of student learning and what support they might need by using not only DC CAS but also other assessments. As it was, there were school-wide activities that aimed at improving student academic performance, but they were not narrowly focused on the state test. The teachers viewed art integration as Jefferson Elementary School‘s initiative.

As the school was not in improvement status, students at Jefferson Elementary School did not receive the letter of invitation to Saturday School. Field trips were not cancelled at Jefferson Elementary School during one month before the DC CAS, while they were cancelled at Central Elementary School. Jefferson Elementary School made AYP in 2007-08. Jefferson Elementary School tended to be buffered from the institutional logics originating from the district.

8.3 CENTRAL ELEMENTARY SCHOOL

Central Elementary School had rarely made AYP over the years. It was one of the 9 transformation schools during 2000-01 because the school's performance on SAT 9 had been consistently low. Teaching staff had to re-apply for their positions. As a transformation school, Central Elementary School implemented a reform program from America's Choice for two years in the early 2000s. All teachers I interviewed for my study came to the school after the restructuring.

Central Elementary School had made AYP only one time, in 2004-05, the final year of SAT 9. Since then, the school had not made AYP, and it was at year one of improvement status (the second year of not making AYP) when I interviewed teachers in 2007-08. The school made AYP in spring 2008 for math through "safe harbor," but it did not make AYP for reading. According to DCPS' guidance, schools that could reduce the percentage of students who were scoring non-proficient by 10% from the previous year would be considered to be making AYP through safe harbor.

Compared to Jefferson Elementary School, Central Elementary School received more policy pressures from the district. Central Elementary School did not have a history of student-centered instruction, nor did it have the resources for subject specialists to support teachers in incorporating district guidance. Compared to Jefferson Elementary School, teachers at Central Elementary School were isolated from other teachers.

The school was located in one of the poorest sections of the city in a neighborhood that was experiencing re-development because of the housing bubble in the downtown and surrounding areas. There were construction projects for a shopping center and townhouses that were priced lower than the houses in other areas of the city. About 80% of students received free or reduced-priced lunch, and almost all students were African American.

The school was an old three-story building with a traditional floor plan. Along the hallway, there were classrooms on both sides. Each heavy metal door to the classroom had a small window. If a visitor wants to know what is going on in a classroom, they must peek into the room through the small window. About one third of classrooms were empty. These rooms were used for after school programs. The school was losing students partly because students had moved to charter schools.

The school did not have as many non-reading and non-math teachers as Jefferson Elementary School did. The school had one Science Specialist Teacher who taught science once per week for 45 minutes to all grades. The school had a music teacher, but did not have an art teacher. The librarian worked part-time, and students had library time every other week. The PE teacher had left early in the school year; hence, students did not have PE class anymore. Teachers had 45 minutes of planning time during these

special subjects, science, library, PE, and music; however, by the middle of the year, the special subjects were offered only twice (music and science). Staffing was not stable. A new principal was assigned to Central Elementary School only a couple of weeks before the school year. A sixth grade classroom teacher left during the school year, so students were split into three groups. One group went to another sixth grade classroom and the remaining two groups of sixth graders went to fifth grade classrooms for about two months until a new teacher was hired.

Central Elementary School allocated longer time on classroom instruction than Jefferson Elementary School did. This school did not have a morning break. Students stayed with their homeroom class until lunch. Students had a 30-minute recess after lunch. The principal decided whether students could go outside and play in the schoolyard or whether they had to spend time in the school auditorium based on their behavior in the cafeteria.

Parent involvement was low. For example, one teacher mentioned that she had only two parents visit her during Parent-Teacher Conference Day.[21] This teacher also told me that grandparents were often raising the students because their parents were divorced or too young. The teachers talked about the lack of parental support as one of the main challenges for teaching.

Students‘ work was posted on bulletin boards in the hallway close to each teacher‘s classroom. When teachers posted students‘ work, which was often a one-page essay or worksheet, they posted them with standards, such as 4.NSO-C.14 Demonstrate

[21] Parent Teacher Conference Day was scheduled four times a year. Teachers were available from noon to 7 pm to accommodate parents‘ work schedule.

an understanding of and the ability to use conventional algorithms for the addition and subtraction of multi-digit whole numbers." Only select student work, usually five to seven portrayals, were posted on each bulletin board.

During the year 2007-08, the school had more activities that directly related to improving DC CAS scores than in previous years. The new principal told staff members that improving test scores was her priority. In January, the school began to post the results of the benchmark tests, DC BAS, and the previous year's DC CAS results, in the hallway. The sign says, "Data Guide Us," and bar graphs showed how many students were at each performance level, i.e., below basic, basic, proficient and advanced by grade level, academic subject (math, reading) and content area, such as geometry, data analysis, and so on. As new DC BAS results came back, the school updated the posts to show the improvement. The information these graphs provided was the number of students who were performing or not performing in each performance category and the number of students who moved to each performance level.

In January, Central Elementary School introduced test preparation sessions, which the district central office suggested to the principals in the district. The principal and counselor proposed a schedule for test preparation. From 11:05 AM to 11:35 AM, all classroom teachers were required to teach test preparation and special subject teachers went to the classroom to assist them. The calendar, created by the principal and counselor, indicated the topic each test preparation session should cover. The principal and counselor went over DC BAS results and listed areas in which their students had performed poorly. For example, one day was designated to reviewing the relationship between numbers and operation, and another day was designated to patterns and data

analysis. The teachers individually prepared materials, such as worksheets or problems from a practice test booklet for their classroom. The school continued this test preparation session until the end of April when the DC CAS was administered.

As Central Elementary School was in improvement status, it received more policy pressures from the district than Jefferson Elementary School. Teachers at Central Elementary School were required to attend professional development sessions organized by the district's central office at the beginning of each semester. The teachers at Central Elementary School attended the district's professional development at the beginning of school year and during the first week of January. Students at Central Elementary School received an invitation to Saturday School. Some teachers at Central Elementary School were recruited to teach at Saturday School from January until the end of April. Central Elementary School did not have instructional initiatives originating from the school, as it was not an incentive school. Rather, the school followed the district's guidance of how the school should prepare students for DC CAS. At Central Elementary School, field trips were cancelled because the district decided to cancel non-instructional activities in March to focus more on test preparation.

8.4 SUMMARY

The difference in organizational contexts between Central Elementary School and Jefferson Elementary School were largely a matter of their exposure to policy pressures

from district and existing instructional contexts for math and science. While Jefferson Elementary School had promoted student-centered, reform-oriented instruction through utilizing Specialist teachers, Central Elementary School did not have these types of resources; thus, individual teachers tended to decide their own methods of instruction. In the next chapters, I present how these differences between Jefferson Elementary School and Central Elementary School shaped the way teachers came across and responded to policy pressures associated with DC CAS.

9.0 FINDINGS: CHANGE OF MATH INSTRUCTION

9.1 INTRODUCTION

In this section, I present how teachers changed their math instruction. I collected data from a total of six teachers who were teaching math. All six teachers experienced the implementation of the new standards and DC CAS, and they were teaching grade levels that took DC CAS during 2007-08. The teachers responded differently to policy pressures associated with DC CAS. In this section, I first present a detailed description of three teachers, Mrs. Green and Mrs. Sommers at Central Elementary School and Mrs. Canright at Jefferson Elementary School. I selected these three teachers because they were different in their views of good math instruction and in their encounters with policy pressures. Existing studies found that teachers' prior views of good instruction shape the way they understand and respond to policy pressures. These three teachers are different from the other three teachers in the following ways. Mrs. Canright's existing instructional practices and views of good instruction were the closet to the reform-oriented math instruction among the six teachers. Mrs. Sommers's existing instructional practice and belief of good instruction was the closet to traditional instruction. Mrs. Sommers and Mrs. Green were at Central Elementary School where they received more policy pressures than the rest of the four teachers who were at Jefferson Elementary

School. Among the four teachers at Jefferson Elementary School, Mrs. Canright most frequently worked with a Math Specialist. Mrs. Canright was the teacher who had the most support, while Mrs. Green and Mrs. Sommers received the least instructional support. The three teachers articulated different understandings of the policy messages associated with DC CAS. By describing these three teachers, I offer the concrete forms of policy pressures, and the context in which these teachers came across policy pressures, to show the data my analysis in the second part of this chapter is based on. For each teacher, I present their range of instructional change since they started teaching to show how their instructional change in response to DC CAS differed from their instructional change during other policy pressures, such as curriculum and standards.

In the second part of this chapter, I present the instructional change identified from the six teachers. By comparing teachers in the same schools and by comparing teachers with different instructional approaches, I describe how these differences shaped their responses to policy pressures. The analysis utilizes within and cross-teacher analysis around the nature of policy pressure, organizational context, and existing instructional practice.

9.2 THREE TEACHERS' CHANGE OF INSTRUCTION

In this section, I present three teachers‘ instructional change from the beginning of their careers to 2008. Three teachers, Mrs. Sommers, Mrs. Green, and Mrs. Canright all had

different views of good math instruction. As existing studies found, their views of good instruction shaped their understanding of policy messages associated with instruction. This was the case for the policy pressures associated with assessment. Their view of how students should learn math shaped their understanding of how they should prepare students for the math test.

9.2.1 Mrs. Sommers at Central Elementary School

I present Mrs. Sommers as a case to show how a teacher who took a teacher-centered approach understood policy messages carried by DC CAS. Mrs. Sommers assimilated most of the policy pressures associated with DC CAS because she saw instructional messages carried by the policy pressures as being in line with her views and practices. Her view of good math instruction and her instructional practice did not change in response to DC CAS as she did not see the difference between her belief about good instruction and good instruction promoted by the policy pressures associated with the standards and assessments. She taught math through teacher-centered instruction, and she focused on mastery of skills through making students follow specific processes to solve problems. During 2007-08, she received more frequent regulative institutional pressure that indicated she had to improve student academic performance than teachers at Jefferson Elementary School. However, she ultimately rejected one policy pressure associated with DC CAS that required her to change her instruction. This finding suggests that Mrs. Sommers‘s belief about good instruction was strong enough to reject a policy pressure that was backed up by regulative pressure.

Mrs. Sommers came into DCPS in the fall of 2003[22] after teaching in other school systems. She graduated with an Education degree during the 1970s. She taught in a testing grade once when the test was SAT9 in DCPS. Then she taught in a non-tested grade, 2nd grade. The school year 2007-08 was the first year she taught in a tested grade of DC CAS.

Mrs. Sommers believes that a teacher should use a specific strategy to teach specific content and skills. In early grades, students learn mathematical concepts through using manipulatives, which she calls a "discovery approach." She believes that students learn the concept of multiplication and division through moving tiles and other manipulatives. Students also need to learn "the process approach," which is groupings of numbers to create new place values; this process approach is used for subtraction and addition of large numbers. Her definition of instructional strategy is a specific method of teaching, such as using manipulatives or teaching place values. Throughout the interview, her view of instruction was "teaching" to students, rather than allowing students to discuss math concepts.

Mrs. Sommers reported many of her students were not at grade level, which was the challenge she had. Mrs. Sommers's 4th grade classroom had approximately 15 students. This class size was common at Central Elementary School. About half of the students were new to Central Elementary School this year, and Mrs. Sommers did not have academic records for these students. Mrs. Sommers told me that she suspected that a couple of these new students would need to be evaluated for special education.

---

[22] DCPS decides the final enrollment for the academic year in October, and teachers are re-assigned to schools based on the final enrollment. The system reassigns teachers who are low in seniority.

Mrs. Sommers's classroom was isolated, and her classroom had a traditional layout. The classroom was on the third floor of the building. A thick metal door blocks sounds from the hallway and other rooms; hence, she rarely heard activities of other classrooms. Desks were lined up, and students worked individually. There were three Apple computers set up at the back of the classroom, but students rarely used them. Mrs. Sommers said she had allowed students to play games when they had finished assigned work, but it was very rare. Mrs. Sommers typically taught skills building for the first 15 to 30 minutes in the morning, moved to reading, and then to mathematics. But this schedule was not strict. She sometimes spent all morning on reading and moved to math in the afternoon.

In line with the district-wide initiative to implement new standards, Mrs. Sommers attended professional development on standards and Everyday Math in 2005-06. She learned about power standards, did activities to unwrap standards, and created lesson plans. She continued attending Everyday Math professional development voluntarily. Among the six teachers I interviewed, Mrs. Sommers most frequently attended district supported professional development.

Although Mrs. Sommers had attended a series of professional development sessions, Mrs. Sommers did not view that the standards and curriculum had changed her instruction. "They might be stated differently but the skills are there. There may be new approaches, like lattice methods, but still you are attacking the same old skills" (Interview on April 3, 2008). Thus, Mrs. Sommers viewed curriculum and standards as

lists of skills and concepts to be taught, and she did not believe that they required different instructional approaches. Although she attended professional development that asked teachers to identify standards and create engaging scenarios, Mrs. Sommers had not seen these activities as promoting a new approach to math instruction that was different from her existing practices. Mrs. Sommers had not seen *Everyday Mathematics* training as promoting new instruction, either. She remembered learning about manipulatives at the training of *Everyday Mathematics*; however, she reported that it had not been new to her because she had been using manipulatives for a long time.

Mrs. Sommers's understanding of the difference between Everyday Math and the previous textbook indicates that Mrs. Sommers viewed school math as a set of skills that students should master and that students learn math mostly through memorization and practice. She saw *Everyday Mathematics* as covering more skills than the previous textbook. She felt Everyday Math was not focused because each unit covers several math skills at one time.

> –Personally, [Everyday Math] is the same (as the old curriculum). Although Everyday Math, they have this kind of approaches we are aiming. In the Math Journal, let's say, you just do not deal with certain skills that you teach to the kids. They also go back to previous skills, for reinforcement, for recall, or for review, so that they do not forget about the past lessons they had...I have to go back to subtraction, addition, grouping, number naming because it is in here (Math Journal)" (Interview on December 5, 2007) .

She says that she only reviewed the very recent lesson when she taught with the old textbook. But with *Everyday Mathematics*, she described that she had to review

—three or four lessons they had for reinforcement. You have to go back and back again to review lessons, review skills. Unless they are able to do this" (Interview on December 7, 2007).

Mrs. Sommers believes that students should learn specific skills in isolation, and she designs her instruction around skills students should master. This approach was different from the approach the district's leadership emphasized in 2005-06. She believed her students were not mastering skills to work on the lesson Everyday Math requires; hence, she often supplemented Everyday Math with worksheets that provided students with practice of specific math skills. For example, she gave students a worksheet that was titled —multiplication," and the worksheet contained computation problems (expressions) and a couple of word problems. The word problems asked simple tasks, for example, —Jamie is planning to have two birthday parties. Find out more about it by multiplying the numbers. She got blocks and has 13 pieces for each box, and she got three boxes of blocks. How many total?" [23] (Interview, January 18, 2008)

Mrs. Sommers's approach to math instruction was consistent even when she used *Everyday Mathematics*. When she taught *Everyday Mathematics* by using Math Journal, she explained each step, rather than making students work on the problems. For example, when she taught a section in which students drew graphs based on the data they collected, she gave students specific guidelines, such as write this number on this line, next write this number on this line.

---

[23] Following is my fieldnote (January 18, 2008): —One student close to me asked: S: Where is the two numbers? And he just started to multiply. (Note: Apparently, he was not reading the problem.) He immediately multiplied two numbers he saw in the problem, and he wrote answer, 19 (Note: It is supposed to be 39, 13 X 3 =39). I asked him why he did it this way, and he pointed out the multiplication table to me, saying 3 times 3 is 9.

When the district introduced the new standards, the district also introduced performance assessment so that the teachers could identify where their students were in the different levels of student performance. Although Mrs. Sommers received training on creating performance tasks and performance assessment at the professional development for the standards, she had not used them. She used worksheets for assessment and grading. She copied items from the report card to her grading book, and she graded how much students mastered each skill or content based on quizzes, which are the worksheets. When she found many students had not mastered skills, she taught again by using a worksheet. She believed that if students worked hard, all students would be able to achieve proficiency as required by NCLB because the students were taught for the standards.

The school year 2007-08 was the first year Mrs. Sommers was teaching in a testing grade of DC CAS. Mrs. Sommers worked with a counselor at Central Elementary School to prepare students for DC CAS. As Central Elementary School was an improvement status school, Mrs. Sommers came across policy pressures that were targeted at improvement status schools. The policy pressures are: attending mandatory professional development, teaching at Saturday Schools, and providing test preparation every day. In addition, Mrs. Sommers voluntarily attended professional development. Mrs. Sommers reported that all these policy pressures helped her to know how to prepare her students for DC CAS.

The district administered the benchmark test DC BAS four times each year. Mrs. Sommers learned how to use DC BAS results from the counselor, and Mrs. Sommers

incorporated information from DC BAS into her instruction throughout the year. The information she found and incorporated into her instruction was the content she should teach. When the first results came back in late November, Mrs. Sommers and the counselor went over the results. Mrs. Sommers said she thought DC BAS was helpful because the information is highly itemized. Teacher Report of the results of DC BAS included the strands missed by individual students. The report also indicated which problem on the test measured particular standards. Every time the results came back to Central Elementary School, the counselor handed Mrs. Sommers a print-out of the results. The counselor wrote the number of the students Mrs. Sommers and another 4th grade teacher needed to move up to the proficiency category for Central Elementary School to make AYP. For example, in February, they needed 10 students to be at proficient. Mrs. Sommers said she had been aligning her instruction and that the DC BAS results helped her ~~–to~~ know where to emphasize" so that she can ~~–g~~ear towards" the standards students missed (Interview on November 29, 2007). She also thinks the report is helpful because ~~–W~~e could gear our questions to the kind of questions that they have [on the test], so they are more or less given the questions here and become familiar with how questions are formulated, so they can attack how to analyze questions. So more or less, they are given tips, how to find the exact answer to the question" (Interview on November 10, 2008).

At the end of December 2007, DCPS announced that it would start Saturday School, and the district began recruiting teachers to teach at Saturday School. Students at Central Elementary School were to go to Saturday School held at a neighboring school that was not making AYP either. Mrs. Sommers began to teach 5th graders at Saturday

School. The district selected a textbook, and the teachers were given a teaching schedule that specified which unit was to be taught each week. She said, "it is just like everyday teaching" (Interview on February 14, 2008), and she hopes the district would adopt this textbook because it is clearer and more on target. She said that the questions in the textbooks were formulated similar to DC CAS and DC BAS.

In early January 2008, Mrs. Sommers attended a three-day district-wide professional development where the teachers received training on DC CAS. Teachers were given a blue print of DC CAS. In one session, the district officials went over the blue print and informed teachers how many questions would appear for each standard and how many points would be given for each question. Mrs. Sommers thought she had gotten a better idea of what would appear on the test, and she could focus more on the instruction on the topics that appear on the test. At school, she learned from the counselor how to use an on-line version of DC BAS Teacher Report. This software allowed teachers to create a worksheet with practice problems that measure specific standards, so teachers could use them for their students. She reported, "everything we do here is aligned to the standards" and "we have to choose only the questions testing those skills" (Interview on January 11, 2008). As Mrs. Sommers was introduced to DC CAS associated materials, she changed her instruction around the content that would be tested but by using instructional methods that she had been using for a long time. The instruction that teaches tested standards dominates her instruction.

Central Elementary School began school-wide test preparation in January 2008. Mrs. Sommers followed the calendar given by the principal and counselor. For example, if the calendar said "addition," she prepared a worksheet on addition for the 4th grade

level. Test preparation time falls after reading class, which Mrs. Sommers has from 9:00 AM to 11:00 AM, although the reading lesson can last until after 11:00 AM. A teacher who does not have homeroom class comes into the room at 11:05. This is the arrangement the school made so that all teachers would be in the classroom to support test preparation. Mrs. Sommers's partner teacher for test prep oversaw student discipline.

The day I observed the test preparation went this way: At 11:05, Mrs. Sommers told students to put away their work and said, "according to the schedule given to us, we are going to do symmetry." She distributed a worksheet. The sheet had diagrams of two-dimensional shapes, rectangular, triangle, heart shape, or lamp shape. According to the instructions of the worksheet, students were supposed to draw a line of symmetry. Mrs. Sommers told students that everything on this sheet is about symmetry. Then she reviewed symmetry with the class by drawing the diagram and line. As Mrs. Sommers explained to the class, another teacher made sure students were quiet and listened to Mrs. Sommers. Then, Mrs. Sommers told students to work on the worksheet (Fieldnote on January 16, 2008). In this test preparation and other test preparations, the tasks Mrs. Sommers presented focused on mastering procedures, and Mrs. Sommers explained the procedure step by step, the same instructional approach she took for regular lessons.

This test preparation often went beyond 11:35 AM, and Mrs. Sommers told me that it was fine because students were learning. The assisting teacher leaves at 11:35, and Mrs. Sommers continues with test preparation until 12:30, lunch time, when the students leave the room for the cafeteria. Mrs. Sommers assimilates test preparation to the degree that test preparation becomes the main math lesson. Mrs. Sommers did not believe that the test preparation had changed her instruction. Test preparation that focused on

mastery of particular math skills through direct instruction was in line with her existing practices.

When Mrs. Sommers taught, she familiarized students with vocabularies so that the students would be able to identify whether they should add, subtract, multiply, or divide. "The vocabulary in math is important…vocabularies used in the word problems. They know where they are supposed to use addition or subtraction, multiplication because that is...what they are weak in. There are so many questions on which …addition is being formulated. So with the kind of vocabularies that are given in the questions, ah this word means-- I have to add. This means I have to subtract…" (Interview on January 10, 2008). Mrs. Sommers also taught shortcuts to solve problems and encouraged students to use the faster way. She told me that some students get stuck on a problem and they run out of time, which would lower their scores.

Mrs. Sommers voluntarily attended professional development that was intended to teach teachers how to prepare for the test. Mrs. Sommers said constructed response was one area she had not known much about. Although the district provided the mandatory professional development for constructed response in reading, the district only provided voluntary based professional development on constructed response for math. Mrs. Sommers reported that she had been trying to learn about constructed response, and Mrs. Sommers attended this voluntary professional development. Although Mrs. Sommers was interested in learning about constructed response, and she knew that students should write in order to improve their test scores, she decided not to spend time on the constructed response. The reason she provided suggests that her views of good instruction shaped this response. Mrs. Sommers said, "these (Note: skills tested in the

multiple choice questions) are the skills they need to cover, and they need to do more on objective choice problems. If they do not know the skills, they cannot do [constructed response problems]" (Interview on April 3, 2008). Hence, she rejected the policy pressures from the district.

Mrs. Sommers's approach to math instruction did not change throughout the year I collected the data. After DC CAS, Mrs. Sommers did math games as the school year was ending. I observed one afternoon class in which she had students playing bingo. Mrs. Sommers reads a problem, which is multiplication and division, such as —three times four equals?" The students raised their hands and answered, and the student who could answer correctly first could cross out the number on his/her own board. She gave rewards, a bag of snacks or stationary, to the first two students who got bingo. As with her regular math lesson and test preparation, the students' tasks were to perform simple computations.

In its current form, Mrs. Sommers's view of good math instruction is centered upon making sure students can compute and get correct answers. Her view that students should learn by mastering computation skills and then move to explaining their thinking had not changed despite the fact that she experienced intense institutional pressures of good instruction promoted by the standards and regulative institutional pressures associated with DC CAS, which required her to make students explain the procedure of their thinking. Among policy pressures associated with DC CAS, she incorporated the message that she had to teach for specific standards. She did not notice that DC CAS' constructed response could challenge her view of how students should learn math. Mrs.

Sommers‘ case shows that despite regulative policy pressures, it was difficult to change teachers‘ views of *how* students should learn math.

9.2.2 Mrs. Green at Central Elementary School

Mrs. Green at Central Elementary School presents a case of how strong policy pressures associated with DC CAS made her focus on topics and skills that were to be tested despite the fact that she had tried to change her instruction to be reform-oriented. Although Mrs. Green viewed her instruction as having changed as she implemented *Everyday Mathematics*, her actual instruction was teacher-centered instruction. Mrs. Green‘s case shows how policy pressures associated with DC CAS led her to reinforce teacher-centered instruction. As with Mrs. Sommers, Mrs. Green‘s instructional practices did not change, but her content of instruction did change.

Mrs. Green began teaching during the school year 2003-2004. DCPS was the first school system she taught in. After teaching different grade levels every couple months during her first year, she had been teaching 5th grade since the fall of 2004. It was her 5th year of teaching and her 4th year of teaching 5th grade when I collected data from her in 2007-2008. Hence, she had experienced teaching in a testing grade for both SAT9 and DC CAS.

Similar to Mrs. Sommers‘s classroom, Mrs. Green‘s classroom is completely separated from other classrooms. Her classroom layout differed from Mrs. Green‘s classroom. Mrs. Green‘s classroom had about 20 students and students‘ desks were put

together so that three to five students sat together[24]. Mrs. Green reported that she sometimes made her students work in groups.

Mrs. Green's math instruction was a combination of student-centered approach and teacher-centered direct instruction in 2008-09. She reported that her educational philosophy, which she calls "constructivist," "guides her instruction." She reported that she had not taught math from this approach during her early years of teaching, although she did think it was a good approach when she had learned it in her education program. She told me her instruction had changed when the district had implemented *Everyday Mathematics*.

> "Honestly, I changed because of the curriculum...First, I was apprehensive to use it because I was like this is stupid, so much a waste of time, kids won't understand it. That was immaturity on my part as a teacher. I did not understand inquiry based research based math lessons. So, I was kind of really apprehensive to use it. But, after reading it and doing it with kids, I realized this is a much better strategy. Teaching them to be more analytical, …develop their higher order thinking skills. But what caused me to start using it is because of the curriculum change. The curriculum changed so I had to change" (Interview on December 11, 2007).

---

[24] [24]. Her classroom size was larger than Mrs. Sommers' class because some 6th grade students were sent to her class after a classroom teacher resigned.

Mrs.Green also reported that around that time she had established her classroom management, which contributed to her using inquiry based lessons. She reported that having strong classroom management is important because students have to work in groups. She also believed that reflecting on her own teaching and learning about math contributed to her re-thinking of how she would teach math. She described how her reflection made her think about changing instruction as follows.

> –Every summer, I usually sit down and…reassess the year, and how I am going to do things differently. So, one year, I sat down and thought about how I understood math because in the past of couple years, students have not performed well on math. It was important that they perform well on the test. I just remember I did not really understand math well until my first year of college. The teacher was really good, and the instruction was great. He basically helped us understand the concepts. He taught us …using that inquiry based methods. So, I thought about the way I learned math. I truly learn math. Okay, I said, I need to start using these things to teach" (Interview on December 11, 2007).

Mrs. Green noticed *Everyday Mathematics* took different instructional approaches from the previous curriculum. She viewed *Everyday Mathematics* as –constructivist" in which –students construct knowledge, and they view concepts based on their own exploration." In contrast, the –old curriculum was more like a tabular model approach. You just taught the math, so it is not a holistic approach. Okay, listen, this is math, use it. So that is different." She reported her instructional change as follows.

> –Before, I probably said, this is triangle...is just this degree, and pentagon is… But now...I have them in small groups. One group had to draw a pentagon and

other group had to draw a quadrilateral. They have to find some of the angles, each of the shapes, and from there we came up with the idea that okay, a triangle has 180, a quadrilateral has 360… So it is more research and inquiry based, instead of me saying okay, this has this. They find this out for themselves" (Interview on December 11, 2007).

Mrs. Green thinks this way ~~–is~~ better for them" because students can ~~–s~~ee that it works by themselves to help them to connect abstract concept and concrete examples…They are not just doing memorization, but actually understand what a triangle is, degree of angle" (Interview on December 11, 2007).

Mrs. Green also reported that she had received some instructional suggestions from another teacher who had been teaching the same grade level. For example, she saw this teacher when the teacher was teaching fractions through making students cut polygons into triangles. Mrs. Green reported that she had seen her students' test scores improve from the previous year since she began teaching this way.

Changing instruction this way was not easy. Mrs. Green mentioned that she had to spend more time on planning and understanding concepts. Furthermore, she found the teacher's manual for Everyday Math did not make it easy to understand what she was supposed to do for the next lesson. She had to spend a long time planning lessons. The teacher's guide for the previous curriculum was clear about what she needed to do.

Although Mrs. Green reported *Everyday Mathematics* had changed her math instruction to be reform-oriented, Mrs. Green's actual instruction was a mix of reform-oriented instruction and traditional instruction. When I asked how students learn math, she answered that they learn through repetition. When I observed her lesson, she began

the lesson with a word problem that asked students to calculate the percentage of savings. Immediately after presenting the problem, her interaction with students was her asking guiding questions so that students follow a particular procedure to solve problems, rather than allowing her students to discuss strategies. In addition, her use of manipulatives rarely allowed students to explore various strategies to solve problems. Rather, she used it to show the results. For example, one lesson I observed was students using tiles to discover how to add and subtract fractions. She distributed tiles to the students, and she used only one question, and pointed the tile to show the back of the tile that indicates ½, and then she quickly moved to whole group instruction. These observations suggest that her instruction for conceptual understanding may be limited to a couple of lesson plans. Hence, it seems her math instruction includes some activities where students work on complex tasks; however, her instruction does not seem to make students explore concepts deeply or ask students to try out challenging and different strategies.

Mrs. Green was aware of her challenges of teaching reform-oriented math, and she viewed her instruction as a mix of reform-oriented instruction and traditional, teacher-centered instruction. Mrs. Green reported that although she strongly believed in the ―constructivist" approach to math and she thought *Everyday Mathematics* was a good curriculum to teach math in this approach, she also reported that she had to go back to direct instruction because she found her students had not mastered skills and concepts they learned in previous grades. She reported her students' lack of requisite skills made teaching Everyday Math difficult for her students as follows:

> ―Everyday Math is a good program but based on a lot of assumptions. Like a lot of things that children are expected to do is built upon what they should have

learned in previous years, but they have not mastered. And it is like you kind of go back and re-teach what they should know to be able to do the activities…. There is a review page called Math Box, and we end up teaching Math Box even though they are not supposed to" (Interview on March 3, 2008).

For example, at one lesson I observed, the original goal was to understand the relationship between fraction, ratio, and percentage. However, Mrs. Green's focus of instruction shifted, making students master the procedure of simplifying fractions. She pointed out a chart that described the procedure of computation. She repeatedly asked the question, "What to do next?" until students became fluent in computation. Although she started the lesson by using a real life scenario, it was narrowed to mastering the procedure of simplifying the fraction.

When the new standards were introduced during the school year 2005-2006, Mrs. Green attended a series of professional development sessions. She learned how to identify power standards and how to unwrap the power standards. She used them in planning instruction. She described her understanding of power standard as follows.

> "Power standards help you to see how you should spend your time. So, if it is not a power standard, if I do not have 90% proficiency, it is okay to move on. But if it is power standards, I need to make sure I have at least 80% proficiency….Power standard is the standard you would see from K through 12….So, by the time they get to me, they should know. For example, we are doing fractions…by the time they get to me, they understand what the numerator

and denominator is…They might not know how to simplify the fraction, but they should know the concept behind the fraction[25]" (Interview on March 3, 2008).

From her view, power standards (identified by the district after 2006) and secured standards (identified by *Everyday Mathematics*) are interchangeable and DC CAS tests those standards. "I basically [focus on] standards that they have to. There are standards in Everyday Math that need to be secured by the time they leave 5th grade, so I spend more time focusing on the standards that need to be secured, and less time on developing standards or introducing standards. Secured standards are the ones that are tested….They (secured standards and power standards) are kind of aligned with each other" (Interview on March 5, 2008).

9.2.2.2 Mrs. Green's Encounter with DC CAS

Mrs. Green reported that pressure to do well on the test had always been high, and she was aware that the assessment had changed her instruction. Mrs. Green reported that pressure to improve student performance on the state test had always been high since Mrs. Green had begun teaching. The district used SAT 9 when Mrs. Green began teaching. She remembers that there was a faculty meeting before the test results were published in the newspaper, and she reported that she felt bad about her students' poor performance.

"You know, the whole school was in the meeting. There was an overhead projector… this is the time when I started teaching. My kids did not do well. My

[25] When I observed her math lesson, she was going over the numerator and denominator because most of the students had not understood which one is the numerator.

kids did very bad. I think it was my first year, and my kids did really bad. That was very discouraging and hurtful as a teacher. But I guess…I mean so much pressure is put on me and principals and schools to perform. And what else are they supposed to do? I do not know, I do not blame the principals or anybody like that….just the general pressure is so much" (Interview on December 11, 2007).

She reported the testing pressure changed her instruction regardless of whether it was SAT 9 or DC CAS. The testing pressure made her cover topics that appear on the test, which sometimes resulted in leaving some students without mastering the content. She described as follows:

~~[~~Before January] I just do not feel comfortable moving unless 70% my students understand (Note: the content or skills)… January, I know they have to be introduced to things because they are all on the test. So I squeeze. I cannot wait until 70% all know it." (Interview on December 11, 2007).

When SAT 9 was administered, as reported in the Washington Post, she stopped regular instruction and taught for the test because the reading curriculum did not cover the grammatical issues that appear on SAT9. Compared to SAT9, she sees DC CAS is in line with curriculum, so it is ~~l~~ess evil" (U.Mrs. Green.IV.071211). However, Mrs. Green did not completely believe that assessment reflected her students' learning. She did not think that standardized tests were sufficient indicators of students' academic performance because ~~s~~ome students do not do well" despite their understanding of academic content (Interview on December 11, 2007). She said in our interview in December that she won't be influenced by the test as much as in previous years. She told me, ~~I~~ feel like this. If I teach, they learn and they do well on test. It may not get

everything, but they will do well. I know this year I am not going to do this as much as I might have (done) my other years" (Interview on December 11, 2007). However, as DC CAS approaches, her regular instruction began to be planned around the standards that would be tested.

Mrs. Green came across policy pressures associated with DC CAS through benchmark testing, professional development, and school-wide test preparation. Mrs. Green reported that there had been more guidance on how to teach for DC CAS during the year 2007-08 than previous years. Mrs. Green reported that she did not receive much guidance about how to deal with DC CAS during school years 2005-06 and 2006-07. In the first year, the district informed teachers about a website that had Massachusetts Assessment so that the teachers could see what the problems would look like. However, she did not remember that the district provided any more specific guidance. The DC BAS started in 2006-2007. Mrs. Green did not pay attention to the results because the results came late in the school year. It was in 2007-08 when Mrs. Green began paying attention to DC BAS around the time the school had a meeting about the test preparation session in January.

In the professional development session conducted by the district in January 2008, she heard that they wouldn't be using power standards anymore, but they should teach for the standards that are tested. She reported as follows:

> I do not use power standards anymore…We were told we do not use power standards anymore. I do try to find standards that are tested and make sure I teach the standards that are tested, and spend time on this. Yes, there are no more power standards" (Interview on March 3, 2008)

But she was not clear on how power standards were different from tested standard and she said the following:

> –they (the district) say there are no more power standards. But I do not know if that is a different form of power standards. So, what I do is, just look at standards, find the ones that are tested and now I focus on those. Yes, I am focused on the test. Everything I teach now is based on the test. So skills I teach are on the test." (Interview on March 3, 2008)

Mrs. Green found –tested standards" on the DC BAS, the blueprint, and the DC CAS students took in previous years (Interview on March 3, 2008). As she arranged her instruction around –power standards" and –secured standards " after January, she planned her lesson to cover –tested standards." She said, –It drives instruction. I group based on the test results, or…regard to what aspect I spend more time, so like if there is a trend like the class did poorly on, let's say, measurement, I spent more time on measurement" (Interview on March 3, 2008).

The school-wide test preparation began in January. Although the principal and the counselor gave teachers a calendar that names the content each classroom should cover, Mrs. Green did not follow this calendar. –Because it is (the plan is) general. I do based on my class, and my class test score is like (She is pointing DC BAS results). So I mean that I take [the calendar] into consideration but I choose the topic that is my specific class. I want to be truthful to my specific class" (Interview on March 5, 2008). As with Mrs. Green, this test preparation lesson sometimes lasts until after 11:30 AM. Students continue with problems on tested standards. In March, the field trip Mrs. Green had planned with a non- profit organization on science and social studies was cancelled.

Although Mrs. Green agreed with good math instruction as student centered conceptual understanding, her actual instruction was close to procedural and teacher-centered direct instruction that focuses on mastering skills, and this instruction took over her regular math instruction. The impact of DC CAS on her instruction was changing the content of instruction.

9.2.3 Mrs. Canright atJefferson Elementary School

I present Mrs. Canright's case to show how a teacher who teaches in line with the approach promoted by the standards and *Everyday Mathematics* incorporated policy pressures associated with DC CAS. Mrs. Canright teaches at Jefferson Elementary School. Similar to Mrs. Green, Mrs. Canright has been teaching in a testing grade since she began teaching. As with Mrs. Green, Mrs. Canright believes that students should learn math through working on complex mathematical tasks. Among the 6 teachers I interviewed, Mrs. Canright's instruction was most closely linked to reform-oriented instruction. Although Mrs. Canright also assimilated the policy pressure associated with DC CAS, Mrs. Canright incorporated different policy messages because Mrs. Canright's understanding of DC CAS was different from Mrs. Green and Mrs. Sommers. While Mrs. Green and Mrs. Sommers understood that they should teach specific content and skills, Mrs. Canright reported that DC CAS was expanding her reform-oriented math instruction.

In this section, I first describe Mrs. Canright's classroom instruction. Then I describe how she responded to the standards and DC CAS. Throughout the section, I present that Mrs. Canright received more support to establish reform-oriented math instruction.

Mrs. Canright began teaching in the fall 2003 through a program that supplies teachers from an alternative route. She was assigned to teach 5th graders at Jefferson Elementary School. For the past two years, she had taught math to all 5th graders and two other 5th grade teachers taught reading and science/social studies.

The school building at Jefferson Elementary School is designed to be an open space. Mrs. Canright's classroom section was at the end of the open floor, and the students' work was displayed on the partitions that provided a sense of separate space. Students' works were displayed on the bulletin boards. I saw individual student work on the topic of data analysis. Each student did a survey on his/her classmates, posing questions to one another such as: What is your hobby? What is your favorite lunch? And, what is your favorite color? The reports (which were written on one-page letter-sized paper) contained numbers, percentages, and ratios. The reports also included a graph and a pie chart based on the data collected.

Mrs. Canright received more training than Mrs. Green at Jefferson Elementary School when she began teaching. Mrs. Canright received on-going trainings for the first two years of teaching between 2003 and 2005 from the program. She met with other teachers in the program and attended workshops every Tuesday after school where she learned classroom management. She reported that although many workshops had not been relevant for her because the program focused on providing teachers to schools in

poor neighborhoods, the best technique she had learned was using a white slate, which allowed her to do a quick assessment during a lesson. The slate allowed students to concentrate on their own tasks; hence, they did not worry about their peers' views. The program evaluated her improvement three times per year. A trainer from the program videotaped Mrs. Canright's lesson, and Mrs. Canright and the trainer discussed possible areas of improvement for her teaching. The program also required her to create a portfolio that included evidence of students' academic growth. As the district did not have an assessment to be used for this purpose, she and other teachers in the program created their own assessment. She believed that this alternative program had taught her the importance of assessment because she had to ask herself how she knew that students were learning. Mrs. Canright stated, "I think the emphasis is not just on bringing kids up, but challenging kids and asking higher level questions, and that was really valuable" (Interview on April 2, 2008).

Mrs. Canright reported that working with a Math Specialist at Jefferson Elementary School helped her learn the discovery approach and individualized instruction. The following account indicates that the Math Specialist regularly worked with Mrs. Canright, and that Mrs. Canright incorporated some of the projects she had learned from the Math Specialist into her classroom instruction. She believed that this approach made students learn. The following account indicated that Mrs. Canright was incorporating math activities the Math Specialist showed in Mrs. Canright's classroom.

> "(The Math Specialist) used to come in every other week. For instance, my first year, he did a lesson with us, where kids trimmed off the verticals of a triangle, and then put them together and then added up to 180. They see they make a

straight line, figuring it out for themselves. So I learned a lot of this discovery-based lessons from him that he does a lot, making children draw their own conclusions based on what they learn. And that way sticks more and students learn better. I did that lesson actually this year. That is something I do not think is in Everyday Math, but I do every year. It is great." (Interview on December 12, 2007).

When Mrs. Canright began teaching, the district used an old curriculum. Mrs. Canright did not think the approach taken by the old curriculum was the best approach to improve students' learning. Mrs. Canright offered the following account about the old curriculum indicating that Mrs. Canright was aware that the old curriculum focused on the mastery of skills, which was different from the instructional approach suggested by *Everyday Mathematics*:

—Everything is like multiplication and division and I thought that it was not interesting for the children. I thought the activities in the teachers' guide in the curriculum require teachers to be creative and go beyond the curriculum. But the worksheets were there, and drills were there and tests were there; and also that requires students to completely master the content before going to the next" (Interview on December 12, 2007).

Similar to Mrs. Green, Mrs. Canright reported that she had established her teaching approach in the third year of her teaching, i.e. 2005-06. Mrs. Canright described the change in her view of how instruction should be structured in the following way:

—I really did not figure out what my approach was until my third year. You come into teaching (Note: with the assumption that) every lesson has an introduction, model practice, guided practice and independent practice. (But) the fact of the matter is that it's not like that everyday" (Interview on December 12, 2007).

Mrs. Canright believes that —making everybody accountable" was the central approach she took to plan her instruction as consistent with the —discovery based approach."

Mrs. Canright's instructional practices and ideas about good instruction are close to reform-oriented instruction. Mrs. Canright assigned activities that included multiple mathematical tasks. For example, in the lesson I observed in March 2008, her class had a month long data collection project called —March Comes like a Lion and Goes out like a Lamb." Students recorded high and low temperatures every day in March. At the beginning of every lesson they reported the temperature, and wrote it on a graph by using the units of Celsius and Fahrenheit. Mrs. Canright quickly asked questions like: How has the temperature changed? How do you think the temperature will change? Is this old saying about March true? She asked students to show how big of a gap they saw in the slope by using their hands. It was about five minutes of activity everyday before they moved to the main topic of the lesson. Students talked about their estimation of the temperature change and described what made them think the old saying was true or untrue. She also included quick computation questions such as difference in temperature, and students wrote the answer on a white slate. The following account indicates that Mrs. Canright believed most students learn better through the discovery approach rather than teacher-centered approaches. In contrast to Mrs. Green, Mrs. Canright thinks that

reform-oriented instruction is for *all* students, particularly students who are performing low.

> –I think that children who learn faster and who learn more easily can learn from copying off the board. For the other 90% of children, especially children with special needs, that abstract way of learning math does not work for them. They need to make connections between what is going on in their hands. We do kinetic things … or we are doing paper fraction things to make that connection. You need a couple of different modes of learning… I also think kids learn best in cooperative groups. I think it is not just pure tutoring but I think by explaining things. Children who are learning faster can further their own understanding, but also their peers can benefit from that." (Interview on December 12, 2007).

Another example of Mrs. Canright's instruction being close to reform-oriented instruction is that she thinks that a math lesson should help students work on multiple concepts and see the connection between the concepts. This view is different from Mrs. Sommers' view that students should learn specific concepts and skills in isolation. Mrs. Canright reported that this approach helps her students to use math concepts in real life and the following account illustrates how Mrs. Canright implemented this approach.

> –If you are teaching fractions in isolation, kids are not able to generalize that to the real world. We just brainstormed and did a whole entry into our notebook. Fractions can be ratios, ratio could be parts of a whole, fractions can be in a number line. I think it helps." (Interview on December 12, 2007).

9.2.3.1 Mrs. Canright's Encounter with Standards and Everyday Math

During the school year 2005-2006, the district introduced the new standards and new curriculum *Everyday Mathematics*. Mrs. Canright attended a series of professional development sessions on the standards and did an activity to identify power standards and the relationship between power standards and other standards. However, Mrs. Canright reported her instruction was more aligned with *Everyday Mathematics* than following the instructional approaches promoted by the standards. Mrs. Canright reported that she had not used the standards-based workbook although she thought learning about unwrapping and power standards was helpful. Mrs. Canright likes *Everyday Mathematics* ~~–b~~ecause it is discovery-based and a lot of (Note: activities) are hands-on."

Although Mrs. Canright liked *Everyday Mathematics*, she did not necessarily follow all of its instructional suggestions. She reported that although the spiral approach of *Everyday Mathematics* made sense for her and she would like to follow this approach, she could not because she had to add concepts and skills that were not required by *Everyday Mathematics* because of the standards, DC BAS and different math curriculum 6th graders use. The 6th graders used Connected Math. She had to align skills and concepts so that students could progress naturally.

9.2.3.2 Mrs. Canright's Encounter with DC CAS

As with Mrs. Green, Mrs. Canright said testing pressure had been high since the very first day she started teaching, and Mrs. Canright reported that the test influenced changes in content. There was SAT 9 when she began to teach. Similar to Mrs. Green, Mrs.

Canright reported that she had changed the order of the content and that she teaches the content that is tested. Mrs. Canright reported that this way of changing her instruction is the same for SAT 9 and for DC CAS. "Reality is, I have to finish ten months of curriculum in seven months to be ready for the testing…Every day we are doing new things, trying to get into all those concepts before the testing." (Interview on December 12, 2007).

Mrs. Canright's view of DC CAS differed from Mrs. Sommers and Mrs. Green in several ways. Mrs. Canright viewed standardized assessment as helpful for evaluating the progress of students who tend to fall behind, but contrary to Mrs. Sommers, Mrs. Canright did not view DC BAS and DC CAS as the primary indicators to measure students' academic improvement. Jefferson Elementary School had been using another assessment system, Star Math and Reading, for many years, which measured individual student's progress in reading and math every four months and showed the grade level of each student. Mrs. Canright and the Math Specialist went over the student reports together and discussed support needs for each student, such as pulling out, individualized instruction, and after school enrichment programs. She believed the NCLB's ultimate goal of making everyone perform at proficient levels on DC CAS was impossible for Jefferson Elementary School because some students came into her class way below grade level.

Similar to Mrs. Green, Mrs. Canright noticed that DC CAS was different from SAT 9; however, Mrs. Canright's understanding of those differences was different from Mrs. Green's understanding. While Mrs. Green noticed the difference between DC CAS and SAT 9 at the content level (meaning DC CAS was in line with curriculum), Mrs.

Canright noticed that DC CAS promotes reform-oriented instruction. Mrs. Canright noticed the nature of the questions was more in line with approaches taken in her regular instruction. The following account illustrates Mrs. Canright's understanding of DC CAS questions.

> —[DC CAS asks questions that were) gravitated to more real life situations, and I think that is a valuable thing to teach in general because no one would ever come and say, do this problem, but you might figure out somehow in the grocery store. That way it has been positive. I am prepping for the test because of that aspect. We have a quiz tomorrow on fractions. Instead of saying, what is the fraction 21 out of 25, I say, if you get 21 out of 25 on this test, what would your score out of 100?…That is the real life application. They stick, also carrying…to generalize mathematical ideas to other situations" (Interview on December 12, 2007).

Mrs. Canright expanded her existing instructional approach by incorporating similar types of questions to the ones asked on DC CAS.

Another policy pressure Mrs. Canright incorporated into her instruction was constructed response, which appeared on DC CAS. Mrs. Canright said a constructed response section where students have to explain how they solved problems on DC CAS had been a positive impact on her instruction. At first, Mrs. Canright did not know how the constructed responses would be scored, although she knew that the points for this section were higher than the multiple choice section. She voluntarily attended a professional development session provided by the district in November 2006. In this training session, a trainer used students' sample responses on a Massachusetts test as the district did not have samples of their own tests. Mrs. Canright received a grading rubric

in the training. After this training, Mrs. Canright showed her students the Massachusetts student sample answers and asked the students to evaluate the sample answers. She used Powerpoint slides used in the training, and she found students were very engaged in pointing out how sample answers did not meet with the perfect answers and in discussing what grade to give sample students. Mrs. Canright reported that since then, every single test and quiz she assigns in her lessons has written response. She reported that she thought it was very important for students to write about math, and writing about math is one thing she had not done before this training and before the testing. She also reported that she felt her students were more ready for DC CAS than SAT 9 because of the written responses the students did. Her understanding of the DC CAS was in line with the intent of math standards that promoted reform-oriented math instruction.

Mrs. Canright did not think DC BAS results were as helpful as Mrs. Green and Mrs. Sommers did because her students had not learned much of the content that appeared on DC BAS. She used the results from another test that Jefferson Elementary School had been using three times a year to measure progress in reading and math to identify students who might need extra help. Hence, as for DC BAS, Mrs. Canright focused on the concepts and skills she had taught, and she went over the concepts and skills that many students had missed.

Mrs. Canright reported that she did test preparation for a week just before the DC CAS, and the instructional approach she used for test preparation was in line with the instructional approach she had used in regular instruction. She asked students to discuss problem solving strategies. In the classroom, she told students to be an –evil test writer" (Interview on April 21, 2008) and come-up with four choices for the answer, i.e., one

right answer, one decoy answer, and two totally wrong answers. I observed students working in groups. First, students had to find the right answer. Then they came-up with three other answers including the decoy answer. In the small groups students talked about why a particular choice could be a tricky decoy answer. At the end, the class shared the answers they created. Mrs. Canright used problems from Massachusetts and other state tests. Mrs. Canright praised students when their decoy answer was the real decoy answer that appeared on the state test. She said she had gotten this idea when she heard from the Math Specialists that multiple choice questions often include two totally wrong answers and one decoy answer. She thought students would be able to come-up with these answers.

Summing up, similar to Mrs. Green and Mrs. Sommers at school Central Elementary School, Mrs. Canright constantly received the policy pressures associated with DC CAS to make students perform better. She re-ordered the content to meet with the test requirements. Furthermore, Mrs. Canright understood policy messages associated with DC CAS as promoting reform-oriented instruction, and she extended her existing reform-oriented instruction by using questions that appeared on DCCAS.

9.3 ANALYSIS

In this section I describe the effect of DC CAS on a total of six math teachers' instructional change. While the previous section presented how three teachers who were different in their views of good instruction responded to policy pressures associated with DC CAS, this section presents the analysis by using concepts from Institutional Theory and Cognitive Approach. I use dimensions like the nature of policy pressures, school context, teachers' existing knowledge, and carriers of policy message to identify what types of instructional change high stakes testing produced or shaped. The analysis suggests that the policy pressures associated with DC CAS did not change teachers' beliefs and approaches to math instruction; thus, the role of DC CAS in changing the core of instruction is minimal. The teachers changed the content of their instruction, but teachers at different schools came across different policy pressures, and the specific school's experience of using assessment in addition to the existence of instructional support shaped teachers' existing instruction. Although the teachers' understanding of how they should change instruction varied, all teachers felt that they had to make students perform well on the test. Thus, policy pressures that were accompanied with the consequence of not complying with the message reached teachers.

I took the following approach to identify how teachers differently responded to policy pressures associated with math high stakes testing. I first identified policy pressures teachers received based on the interviews, observation, and policy documents.

Second, I categorized the content of each policy pressure by deciding if it carried a message about instruction or not. Third, I categorized teachers' responses to the policy pressures by using Coburn's categorization (Coburn, 2004). As Coburn's categorization articulates instructional change in relation to existing instruction, the new instruction in response to policy pressure varied teacher by teacher, although I assigned the same categorization. As for teachers who taught close to reform-oriented instruction, for example, Mrs. Canright, "assimilation" was teaching tested content by making students discuss concepts. For teachers who had been teaching with a more traditional approach, for example, Mrs. Sommers, "assimilation" is teaching tested content by making students follow specific approaches to solve problems. Once I created a list of teachers' responses to policy pressures, I compared individual teachers' responses over time (within teacher analysis) and compared responses across teachers (cross teacher analysis) to identify the trend across six teachers. The table below presents all the policy pressures and responses I analyzed. An outline of this section is as follows. First, I present what types of policy pressures were conveyed to teachers. Second, I present how a school's organizational context shaped the policy messages the teachers came across and how these messages shaped teachers' responses. Third, I present how teachers' existing instructional practices and their views of good instruction shaped their responses to policy pressures.

9.3.1 Nature of Policy Pressure

In this section, I present how the nature of policy pressure, one of the three aspects identified by the Cognitive Approach as shaping teachers' instructional change, shaped the six teachers' responses to policy pressure. My data shows that the policy pressures associated with DC CAS included the following two types of policy messages: a) a message that teachers have to improve student academic performance, and b) a message about how to change classroom instruction. Distinguishing the content of the message was important because while all teachers felt pressure to improve students' academic performance, teachers came across different policy pressures that shaped their understandings of *how* they should change instruction. The policy messages about instruction were likely open to teachers' interpretation because the change of instruction was not monitored, but all teachers knew that they had to make students perform well on the test.

My data confirmed the findings from previous studies on the impact of instructional policy on teachers' classroom instruction. Previous studies on teachers' response to instructional policy identified that teachers tend to incorporate concrete messages and messages that teachers came across frequently (Coburn, 2001, 2005, Spillane & Jennings, 1997). In my study the teachers tended to notice and incorporate policy messages that were clear to them and messages they came across more often. The messages I examined included messages that suggested change of instruction and messages that required teachers to improve students' academic performance. The teachers were likely to incorporate policy messages that suggested changing the content

and order of instruction rather than the messages that suggested different instructional approaches because they rarely noticed the messages that suggested changing their instructional approach. The carriers of policy messages also mattered. In line with existing studies (Coburn, 2001, 2004), teachers tended to incorporate instructional messages that were suggested by colleagues who shared similar instructional approaches and who often interacted with the teachers.

The outline of this section is as follows. First, I present two types of policy messages associated with DC CAS. Second, I describe characteristics of policy messages that reached teachers. Third, I present how the frequency, strength of persuasion, and timeliness of policy pressures shaped teachers‘ responses. Fourth, I describe how the carrier of policy messages shaped their responses.

9.3.1.1 Content of Policy Messages

The teachers came across two types of policy messages associated with high stakes testing. One type of policy message is that schools have to make AYP, which was required by the NCLB. The policy messages in this category included a general message that teachers have to make students perform well on DC CAS and general information about how many students the teacher should move up to proficiency level to make students perform. Another type of policy message was about instruction, such as the proposal that teachers should do test preparation or that instruction should focus on tested standards. There were two types of messages about instruction. One message suggested the content of instruction and another message was how the teachers should teach math or

teach for the test. The table below presents the list of policy pressures teachers came across by the content of message.

Table 9 List of Policy Pressures and Teachers' Responses

Policy pressure associated with	Teacher	Policy Pressure	Carrier	Response	Aspect of instruction influenced by the policy pressure	Response (in relation to existing instructional practice)
SAT9 (2003)	Mrs. Green	Saw the test results at a faculty meeting	Test results	She felt bad, embarrassed, disheartened	No direct impact. Made her reminded that students have to perform well.	N/A
DC CAS DC BAS	Mrs. Green	Talked with colleague about how to prepare students for test.	Colleague	They agreed with each other that they focus on teaching secured standards.	Content Focus on teaching these 'ds.	Assimilation
SAT9 DC CAS	Mrs. Green	Received test results	Test results	She went over with individual students and set goals for next year.	Not specific on instruction. Reinforcing existing approach	N/A
DC CAS(2004-05)	Mrs. Green	Was told that —This will going to be the test, go and look for	Principal. Web page.	She looked at Massachusetts test, and checked what skills are tested.	Content Made sure that those skills are taught.	Assimilation or Parallel Structure

		sample questions (on the web)"				
DC BAS (2006-07)	Mrs. Green	Received test results of DC BAS late	Test results	She did not pay attention to it	No impact on instruction	**12.1.1** N/A
DC CAS, DC BAS (2007-08)	Mrs. Green	Heard about school performance more often than before.	Principal, colleagues, test results,	She felt much pressure on reading and math this year. "The district really wants to see improvement."	No specific change of math or science instruction, but it reinforced her understanding that she had to improve test score. Consequently, science became secondary for her and other teachers	N/A
DC BAS(2007 Nov)	Mrs. Green	Received DC BAS results much quicker than previous year.	Test Results	She did not look at it.	No impact on instruction	N/A
DC CAS, DC BAS (Jan 2008)	Mrs. Green	Principal and data specialist proposed starting test preparation time	Principal, Data specialist, calendar,	She set up test preparation time, but she did not follow the calendar of the content	Schedule of lesson by setting up test prep time.	Assimilation
Jan 2008 DC BAS, DC CAS	Mrs. Green	She was told they won't use power standards anymore, but	PD	She does not understand how they are different, but she teaches for tested standards	Content More focus on tested standards.	Assimilation

		they should teach around the standards tested				
DC CAS DC BAS (March 2008)	Mrs. Green	She was told by the principal that field trip would be cancelled	Principal, (but also this is from district)	She cancelled.	Schedule of lesson changed, science and social studies components were cut.	N/A (change of schedule)
SAT 9 (prior to 2005)	Mrs. Reed	She saw test items, talked with colleague.	Test itself	She started to use terms appear on the test. Because she does not want students to be tested on the topics/skills they do not know.	Content Used terms that were used in the test. Exposed students to the topics that had appeared on the test.	Parallel Structure
DC BAS (2006-07)	Mrs. Reed	She saw DC BAS results handed to her.	DC BAS results	She thought it was very general (such percent of students missed geometry…) so she did not see it helpful.	No change of instruction	N/A
DC BAS (2007-08)	Mrs. Reed	Principal explained what is in teacher report of DC BAS, Grade level teachers met and went over the results and decided to ask students	DC BAS results, PD, Agreement among colleagues	She thought it is very helpful because the teachers know what students are actually missing She first started asking questions to the students, but soon she taught how to solve the problems.	Content She went over the problems with the students, and re-taught the concept to make sure students cover it. She taught the content in a way that fits with her existing instructional practice.	Assimilation

		how they thought about the problem.				
SAT 9 DC CAS	Mrs. Reed	She administering test and saw test problems.	Test	She thinks there is not much difference between SAT 9 and DC CAS regarding how to make her students ready for the test. She makes sure students know the concepts and skills that are tested.	Content She went over the content and skills covered by the test.	Parallel Structure
DC CAS (2005-06)	Mrs. Sommers	She heard about DC CAS from other teachers. She heard constructed response was new.	Colleagues	She did test preparation for the second graders by using SAT 9 materials for test preparation.	Content She did test preparation.	12.2 Parallel Structure or Assimilation
DC CAS (2007-08)	Mrs. Sommers	She went to PD on constructed response on math.	PD	She wanted to learn about constructed response. But as time goes by, she decided not to focus on it because students need to learn skills first.	No change in instruction.	N/A
DC CAS (2007-08)	Mrs. Sommers	She heard a new principal saying his goal is improving student	Talk from principal	She strongly agreed with it.	No change in instruction. It reinforced her view of what is important for students.	N/A

		academic performance				
DC BAS (Dec 2007)	Mrs. Sommers	Data specialist and she went over DC BAS results	Data Specialist, DC BAS results	She thought that she could gear her instruction towards the skills students missed.	Content Helped her to focus on the skills students missed on DC BAS.	Assimilation
DC CAS, (Jan 2008)	Mrs. Sommers	At a district wide PD, they went over how many standards appeared on DC CAS.	Blue Print	She thought it helps her. She thought she had better idea of what would be tested.	Content Helped her to focus on the standards that appear on the test.	Assimilation
DC CAS, DC BAS (Jan 2008)	Mrs. Sommers	Principal and data specialist explained how to use DC BAS results. They proposed starting test preparation	Principal, Data specialist, DC CAS results, Think Link, Calendar	She agrees with it. She set up test preparation time, and she taught the content specified in the calendar. As for using Think Link (creating practice problems), she understood that she has to choose questions that are only on the test.	Schedule and content of lesson. Change of the content of instruction and schedule.	Assimilation
DC CAS (Jan 2008)	Mrs. Sommers	She attended training for Saturday School	PD, Test preparation textbooks	She does not think anything is new, but she thinks the textbook should be in line with the standards tested on DC CAS, so she wants to use it at her class.	Not direct impact on her regular instruction.	N/A
DC CAS	Mrs.	Data	DC BAS results,	She thought she needed a	Not direct impact on her	N/A

DC BAS (Feb 2008)	Sommers	specialist delivered DC BAS results and showed her how many students Mrs. Sommers and another 4th grade teacher need to move up.	Data specialist.	couple more students moved to the proficient level by teaching for the standards tested.	instruction. Reinforced her approach of teaching for the test.	
SAT9 (Prior to 2005)	Mrs. Canright	Administered SAT 9 and saw problems.	Test	She saw the test and found it is mostly computation problems. Students did well on these types of problems.	Content She did test preparation to cover the content, She thought it was stressful.	Parallel Structure or Assimilation
DC CAS (2005-06)	Mrs. Canright	She administered the test, saw problems.	Test	She thought the problems appeared on the test reflects real life problems, and she thought these problems are good.	Content with conceptual understanding. She incorporate that aspect of test preparation	Assimilation
DC BAS	Mrs. Canright	DC BAS Results	Teacher Report	She went over and saw what students missed. She found number sense is weak, so she went over again by incorporating tasks that use number sense into multiple presentations.	Content She re-taught skills and concepts students missed.	Assimilation
DC CAS (2006-07)	Mrs. Canright	She attended PD to learn	PD	She thought it was helpful to decide how she	Content She made her students	Assimilation

		what evaluators would be looking at in constructed response.		should teach and what suggestions she should give to the students for constructed response.	grade hypothetical students‘ answer.	
DC CAS	Mrs. Canright	She received Blue Print from the principal.	Blue Print	She went over Blue Print and decided the content to teach	Content She covered the skills tested on the test.	12.3 Assimil ation
12.4 DC BAS	Mrs. Caracell i	She received DC BAS results.	Test results	She did not see the test results as helpful because many skills tested on the test she did not teach. She looks at results, and re-taught skills students missed.	Content She re-taught the content and skills students missed.	Assimilation or Parallel Structure
DC CAS	Mrs. Caracell i	She administered DC CAS, saw problems and results.	Test and results	She thinks DC CAS is well rounded and there is not much specifically teach for the test	She continued teaching curriculum with one week of test preparation.	Parallel Structure
DC CAS	Mrs. Miller	She heard from someone constructed response is included in DC CAS. But she did not remember having	Test, Someone (somewhere..not specific)	She just did students to give me as much as possible you can get She told students which answer is good, which is not.	Content She did constructed response questions with students.	Parallel Structure or Assimilation

		specific training on it.				
DC BAS 2006-07	Mrs. Miller	Received DC BAS results.	Test results	She did not use it because they did not get specific information. It was very general.	No specific impact on instruction.	N/A
DC BAS (Nov. 2007)	Mrs. Miller	Received DC BAS results	Test results	She saw the results to get an idea of which students are behind, or may need special attention.	No specific impact on instruction. She used information to get an idea of which students she may need to pay more attention to.	N/A
DC BAS 2007-08 Jan	Mrs. Miller	She learned from the principal about what DC BAS results include, and she and other teachers went over the test results.	Test results, PD	She went over the problems with students. She thinks the information is very specific and helpful.	Content She went over the problem with the students.	Parallel Structure or Assimilation

Among the policy messages about instruction, teachers were more likely to come across messages that suggested the content of instruction as opposed to the pedagogy of instruction. This difference seemed to derive from the carriers of the message. While the policy messages about the content of instruction (i.e. what teachers should teach) were carried through documents, for example, Teacher Report of DC BAS results, messages about the pedagogy of instruction were communicated verbally during professional development or informal situations, such as hearing how to teach at professional development sessions. For instance, at Jefferson Elementary School, a Math Specialist suggested to the teachers that they encourage students to write two different ways, one with numbers and another with words so that students will be able to provide a comprehensive answer to the constructed response questions during a professional development session No teacher objected to this suggestion, and no further discussion about how this suggestion might fit into individual teachers' instructional practices happened. Rather, these types of suggestions were mentioned without further discussion. Hence, compared to the policy pressures associated with the content of instruction, policy pressures associated with strategies of instruction were short-lived and weak in persuasion.

Among these three types of policy messages, the policy messages that teachers should make students perform well on the test and policy messages about the content to be covered in the classroom instruction reached teachers. All teachers knew that they had to make students perform well on the test, and they felt they must make students perform well on DC CAS. All teachers reported that they were covering the content that would

appear on the test. The teachers varied if they noticed the message about instructional approach in the policy pressures.

While the district intended that standards, assessment, and curriculum were to send a uniform message about good instruction, none of the six teachers came across policy pressures that explained how standards, curriculum, and DC CAS were related. The teachers heard in the professional development in 2005-2006 that DC CAS was aligned with the power standards. However, they found that this was not the case when the first DC CAS was administered. Since then, the teachers did not come across how these three components were aligned.

9.3.1.2 Characteristics of Policy Message

My data supported the previous findings that teachers tended to incorporate instructional messages that were concrete and that teachers tended to incorporate policy messages that suggest changing the structure of instruction rather than changing pedagogical methods of instruction. The Cognitive Approach suggests that the nature of policy pressures matters in teachers‘ responses, and existing studies found that instructional messages that are concrete tend to reach teachers. (Spillane & Jennings, 1997). The data from the six teachers also indicates that the teachers tended to notice a policy message that was clear. The teachers‘ use of DC BAS results presents this example. While DC BAS was introduced in 2006-2007, the teachers in my study rarely looked at them, and they did not see the information from DC BAS as shaping their instruction until the middle of the 2007-08 school year. The teachers‘ responses to DC BAS results changed after they

received training on how to analyze the test results in January 2008. The district instructed principals to teach their teachers how to read DC BAS results. When I interviewed the teachers earlier in the school year, they were not looking closely at DC BAS results. Mrs. Sommers at Central Elementary School did not see the results. Mrs. Miller, a teacher at Jefferson Elementary School, reported early in the year that she looked at where students were to get a general sense of which students she might need to pay more attention to in the classroom. "Just kind of, to me, [the DC BAS] is to target certain kids, versus to see 'okay, you are in trouble in number operation or measurement.' It is last thing (for me to do) this matching kid with this (specific skills)…I think it is more to target kids (who are) lower in general" (Interview on December 10, 2007) .

After receiving the instruction from the principal on how to look at the data in January 2008, most teachers in my study found the data helpful because the data showed what problems and what standards students missed. Mrs. Miller's account after the training indicates that her view of the usefulness of DC BAS results changed. "We were never given our kids' exact information like that. I never knew before…like…that Denny missed number two and put this, instead of this. So, we did not go over it. Ah, I could have gone over the packet. And it is like a whole entire packet, I would not have known how many kids have missed. This is, you can target… So, patterns across 4th grade… So, this year is much better... It is the best" (Interview on January 8, 2008). As it is, the training helped teachers to see DC BAS results. The teachers viewed the information as concrete and clear. Teachers incorporated the standards students missed on DC BAS.

My study also supports the findings from existing studies (Coburn, 2005; Diamond, 2007) that teachers incorporated the instructional messages that required structural change, such as the change of scheduling, rather than the messages that required changing teaching practice. At Central Elementary School, although both Mrs. Sommers and Mrs. Green introduced the test preparation session from 11:05 AM to 11:35 AM, Mrs. Green did not follow the calendar given by the principal because she thought the calendar did not reflect her students' needs. Rather, she made a plan for her class based on her students' DC BAS results[26].

9.3.1.3 Frequency and strengths of policy pressures

In line with existing studies (Coburn, 2001, 2004), teachers in my study tended to notice policy messages that they came across frequently. The teachers viewed the pressure as strong when they came across similar policy messages. In my study, the teachers followed the policy messages that were carried as requirements; however, these messages did not cover how teachers should teach math in the classroom.

Teachers' responses to policy pressures associated with DC BAS from the school year 2006-07 and school year 2007-08 show that the more frequently the teachers came across the message, the more strongly the teachers felt the importance of the message. During the school year 2006-07, although the district provided Blue Print and DC BAS, the teachers at Central Elementary School did not receive the follow-up messages about

[26] Later, the principal and counselor proposed dividing each class into two groups, and an assisting special subject teacher took care of one group. The teachers pushed back against this proposal with the reason that they were afraid that their students would miss some of the instruction by not being with them, and the special subject teachers might teach with different approaches. Consequently, this proposal did not go anywhere, and the classroom teachers continued teaching all their students.

DC BAS, nor did they receive information about how many students they must move to the next level. During 2007-08, the teachers frequently came across the policy message that they had to improve students‘ performance. The counselor informed Mrs. Green and Mrs. Sommers how many students they must move up to proficiency level; they saw the DC BAS results periodically; they were asked to teach at Saturday School; and they were required to do test preparation. These changes in 2007-08 made Mrs. Green and Mrs. Sommers see that the district was very concerned about improving students‘ academic performance. Mrs. Green and Mrs. Sommers‘ observation about the priority of the district and school led these teachers to continue with the test preparation.

The strength of persuasion influenced teachers‘ responses, but the messages that were mandatory for teachers and the messages teachers viewed as mandatory did not cover instructional approaches. The messages teachers came across required teachers to change the schedule of lessons and to organize specific activities with fellow teachers, such as have a meeting to discuss instruction, the improvement in proficiency numbers, and the content of class instruction. No teacher came across policy messages that teachers must teach math in specific ways in the classroom. For example, although the principal at Central Elementary School required teachers to follow the test preparation calendar, the principal did not provide specific guidance about how teachers should teach for the test. Most teachers followed these requirements. However, as I described in the previous section, the closer the core of instruction, such as what a teacher actually teaches in the classroom, the more latitude the teachers had—such was the case with Mrs. Green not following the test preparation calendar provided by the principal.

9.3.1.4 Carriers of the Policy Messages and Teacher's Response

In line with existing studies on teachers' responses to instructional policy (Spillane & Jennings, 1997; Coburn, 2004) and the study of teachers' perceptions of what influences their classroom instruction (Diamond, 2007), teachers in my study tended to notice and incorporate messages about instruction that were carried by their colleagues or principals. All teachers reported that they learned how to use DC BAS results when they received training from principals.

As Coburn found in her study of collective sense making of reading policy among early grade teachers (Coburn, 2001), the teachers in my study also confirmed their understanding of how they should prepare their students for the test with colleagues who shared the same view of instruction. Mrs. Green saw her partner teacher was knowledgeable about DC BAS, and Mrs. Green confirmed with the partner teacher that their students' biggest problem was not mastering fundamental skills; thus they should focus on teaching these skills. At Jefferson Elementary School, Mrs. Canright picked up an idea from the Math Specialist when she planned test preparation. The Math Specialist taught her that a multiple-choice problem often contains one decoy answer, and Mrs. Canright decided to make students create test problems that included decoy answers. Since Mrs. Canright began teaching at Jefferson Elementary School, the Math Specialist had helped her to introduce reform-oriented instruction.

Another example is the non-response by the teachers, at Jefferson Elementary School, who had not worked together. Teachers at Jefferson Elementary School did not respond to Mrs. Canright's suggestion to look at Power Point slides for teaching constructed responses. The teachers' non-response to Mrs. Canright's suggestion seemed

to derive from the fact that teachers at Jefferson Elementary School often work within the same grade level. In addition, other teachers did not view constructed response in the same way Mrs. Canright did. Mrs. Canright viewed constructed response questions as being in line with her view of good math instruction, and she thought constructed response had brought about positive change in her instruction. Although other teachers at Jefferson Elementary School also thought constructed response questions were important and their students should do constructed response well, they did not view the constructive response as changing their classroom instruction.

Although the colleagues were able to reach teachers, these teachers shared similar views of instructional approach. Thus, the instructional messages carried by colleagues did not change the teachers‘ instructional approaches. Teachers added new content or activities suggested by their colleagues, but these activities and content did not change the teachers‘ instructional approach.

9.3.1.5 Summary

This section presented how the following dimensions associated with the nature of policy pressure mattered to teachers‘ instructional change: a) the content of policy messages, b) the frequency of the encounter with the policy messages and the strength of persuasion of the messages, and c) carriers of the policy messages. While policy messages that teachers had to make students perform well and the messages about the content of instruction reached teachers, teachers rarely came across the policy messages that asked them to change their instructional approach. The teachers were more likely to notice and incorporate instructional suggestions when the messages were concrete, were

conveyed frequently, and were carried by colleagues. The teachers incorporated policy messages that they saw as mandatory and that they saw were carried by school administrators. However, these messages did not cover how teachers should be teaching in the classroom.

9.3.2 Organizational Context

The comparison of the teachers‘ responses between school Jefferson Elementary School and Central Elementary School suggests a school plays a part in shaping teachers‘ responses to DC CAS. In this section, I contend that the school difference contributed to teachers‘ responses to high stakes testing in the following ways: a) by mediating policy pressures originated from the district, b) by shaping instructional practices and a shared norm of how to make students successful; and c) by providing experiences using assessment results to identify instructional support. Below, I first present how teachers at Jefferson Elementary School and Central Elementary School differed in their responses to DC CAS. Then I present the above three aspects as the school differences that contributed to the teachers‘ different responses to high stakes testing.

9.3.2.1Different Responses

The teachers at Jefferson Elementary School and Central Elementary School responded differently to policy pressures associated with DC CAS. Three out of the four teachers at Jefferson Elementary School incorporated testing pressure by taking parallel structure, such as having as 15 minutes a day of test preparation lessons throughout the year or

having a review session for about a week before the test. The teachers at Jefferson Elementary School were aware that these test preparation lessons were separate from their main instruction. Among four teachers at Jefferson Elementary School, Mrs. Canright has been most influenced by the policy pressures associated with DC CAS because she changed the order of instruction and added the content that was not covered by *Everyday Mathematics*, and also changed the format of her assessment to include constructed response.

Policy pressures associated with DC CAS guided regular instruction at Central Elementary School. At Central Elementary School, both Mrs. Sommers and Mrs. Green taught for the tested standards, hence teaching for test was more prevalent at Central Elementary School than at Jefferson Elementary School. Both Mrs. Sommers and Mrs. Green extended test preparation time until lunchtime, which replaced their regular lesson. While teachers at Jefferson Elementary School generally incorporated policy pressures associated with the test as parallel structure, teachers at Central Elementary School incorporated policy pressures through assimilation by planning instruction around the tested standards. DC CAS did not change any of the six teachers‘ instructional approaches. The difference was in time spent and the degree to which the content of instruction was driven by DC CAS.

In the following three sections, I present the differences between Central Elementary School and Jefferson Elementary School. I contend that the school differences shaped the way teachers came across policy pressures and the teachers‘ understanding of those policy messages. The overview of this section is as follows. First, I describe how the schools mediated the policy pressures originating from the districts;

thus, the teachers at the two schools came across different policy messages. Second, I describe the difference in staffing and instructional support, which shaped the existing instruction of math curriculum and shaped beliefs about good instruction among teachers. Third, I describe the experience of using assessment results and how that experience shaped the teachers‘ views of usefulness of DC BAS results, which also shaped the way teachers responded to the DC BAS.

9.3.2.2Mediation of Policy Pressures from the District

The frequency of the policy messages teachers came across and the content of policy messages teachers came across differed between Central Elementary School and Jefferson Elementary School. The teachers at Central Elementary School more frequently came across policy pressures that carried the message that teachers must teach for the tested standards than the teachers at Jefferson Elementary School. The teachers at Central Elementary School were more likely to come across the policy pressure that they must improve student performance on the test. This difference might come from the fact that Central Elementary School was in improvement status and it received policy messages from the district, while Jefferson Elementary School was a high performing school, and it was buffered from the policy pressures from the district.

The teachers at Jefferson Elementary School were buffered from the district‘s policy pressures because the district allowed Jefferson Elementary School to provide school initiated professional development, and the teachers at Jefferson Elementary School were exempted from mandatory professional development held by the district.

Thus, the teachers at Jefferson Elementary School received policy pressures that were mediated by principals or specialist teachers. The difference in the school's improvement status was important because while the teachers at Central Elementary School received the message that they must teach for tested standards, the teachers at Jefferson Elementary School did not come across this specific message to the same degree of intensity and concreteness. Below I present how the school level difference contributed to the these two aspects of teacher experience, i.e., a) exposure to the message that teachers must teach for the tested standards, and b) frequency of policy pressures that teachers must improve student performance on DC CAS.

A school's role in mediating policy messages was evident in how teachers came across the district's guidance on how teachers should use DC BAS results. In January of 2008, teachers in the district, including Mrs. Green and Mrs. Sommers at Central Elementary School, attended a three-day district-wide mandatory professional development. In the professional development, teachers went through the Blue Print that listed the standards that would be tested in the upcoming DC CAS, sample problems, and possible points for each problem. Both Mrs. Sommers and Mrs. Green reported that the message they took away from this professional development was that they must teach for tested standards. In addition, in the full day professional development session on constructed response that Mrs. Sommers attended, I observed that the facilitator emphasized lesson planning that started from tested standards. Soon after school resumed in January 2008, the principal and the counselor at Central Elementary School taught teachers how to analyze DC BAS results. It was a 40-minute session in the morning, and the principal and the counselor talked to a whole group of teachers. Mrs.

Sommers and Mrs. Green reported that they learned that they had to look at the standards students missed and re-teach these content areas. In this session, the principal and the counselor proposed to start a test preparation session everyday, and they gave teachers a calendar that shows skills and content areas the teachers should cover in each session. Mrs. Green and Mrs. Sommers reported that their understanding was that they should teach for the standards that students missed on DC BAS.

Teachers at Jefferson Elementary School received a different message about how to use DC BAS results. Because Jefferson Elementary School had been a high performing school, teachers at Jefferson Elementary School were exempted from the professional development sessions provided by the district. Instead, Jefferson Elementary School held an in-school professional development in which the principal explained how to read the Teacher Report of DC BAS. After explaining information on Teacher Report, the principal asked teachers from each grade level to get together and go over the results of DC BAS to see what the they could tell from the data. For about two hours, the teachers went over the results and talked about what DC BAS covered, what they think about the results, and why their students missed certain problems. In the afternoon, the Math Specialist facilitated professional development in which teachers looked at DC BAS problems and talked about why students might have chosen wrong answers. The meeting resembled a brainstorming session in which teachers talked aloud about their thoughts on how their students might have understood the DC BAS problems. In the meeting, the teachers talked about some strategies they might use, such as making sure that students always write down their understanding on the test sheet, but there was no proposal about what or how teachers should teach in their classrooms. In this case, the

teachers at Jefferson Elementary School did not come across the message that they must teach for tested standards.

The second difference was that the teachers at Central Elementary School more frequently came across the message that teachers must improve students‘ performance on DC CAS than the teachers at Jefferson Elementary School. All teachers in my study reported that they knew the school‘s improvement status and that they knew they must make sure that students performed well on DC CAS. All teachers reported that they come across policy pressures as they hear about the test in the media or talk about it with other teachers. A difference between Jefferson Elementary School and Central Elementary School is that the principal and counselor periodically conveyed this message to the teachers by presenting specific improvement targets. Also teachers at Jefferson Elementary School and Central Elementary School differ in their degree of participation in the professional development that aimed at improving student performance on DC CAS provided by the district. At Jefferson Elementary School, the principal did not provide this specific guidance to teachers, and the teachers did not engage in the activities the district planned to improve test performance. . At Central Elementary School, every time DC BAS results came back, the counselor informed each teacher how many students in each grade level needed to move up. The school posted the results of DC BAS, including the change in the number of students who were at each proficiency level. The principal told teachers that improving the DC CAS scores was the priority. The teachers were asked to teach at Saturday School, and the school cancelled fieldtrips. Teachers at Jefferson Elementary School did not come across policy messages that were this specific.

The school continued with professional development on art integration, and the teachers were not involved in Saturday School.

The comparison of Central Elementary School and Jefferson Elementary School suggests that in addition to a school's status in relation to the district, the administrators also shaped the policy messages teachers received at both schools. The principals mattered because the teachers received different policy messages about what they should do to improve student academic performance. The principals played an important role in setting a framework of what the teachers could potentially understand from the policy messages by facilitating activities that guided the scope of discussion. At Central Elementary School teachers were told to teach for the tested content, while at Jefferson Elementary School teachers were asked to think about what the data revealed. At Central Elementary School, the teachers taught for tested standards, while the teachers at Jefferson Elementary School analyzed students' thinking, and the discussion of instruction was not limited to teaching for the tested standards.

9.3.2.3Existing Curriculum

The school's existing instructional support and scheduling also seemed to have contributed to the different responses from teachers at Jefferson Elementary School and Central Elementary School. The school context shaped the existing instructional practices and beliefs about good instruction among teachers. While teachers at Jefferson Elementary School were more likely to continue to teach *Everyday Mathematics* and incorporate the skills and content that appeared on DC CAS as a parallel structure, the

teachers at Central Elementary School began to plan their instruction around frequently tested standards as they encountered policy pressures associated with DC CAS. This difference seemed to derive from the fact that while Jefferson Elementary School had a Math Specialist who supported teachers to implement *Everyday Mathematics*, teachers at Central Elementary School had not had this type of support. In addition, existing scheduling at Central Elementary School and Jefferson Elementary School made a difference in enforcing implementation of *Everyday Mathematics*.

Most teachers at Jefferson Elementary School closely followed the order and content of *Everyday Mathematics. Everyday Mathematics* is a highly prescribed curriculum, and the district implemented it in the school year 2005-2006 through district-wide professional development. The teachers at Jefferson Elementary School often found that they did not have enough time to cover everything in the teacher‘s manual. Hence they tried to go through *Everyday Mathematics* as much as possible. Although there were some gaps between *Everyday Mathematics* and the district standards in the content to be covered, the teachers at Jefferson Elementary School reported that *Everyday Mathematics* drove their instruction. The teachers added content that appeared on DC CAS, but was not covered by *Everyday Mathematics*.

The teachers at Jefferson Elementary School reported that they liked *Everyday Mathematics* and they thought if they taught *Everyday Mathematics*, students would do fine on DC CAS. For example, the account provided by Mrs. Caracelli, who reported that she closely followed the Math Journal of *Everyday Mathematics*, indicates that she did not see the need to change her instruction to make students ready for DC CAS.

—What I like about this standardized test is that it is very generic and …so they kind of have to know the things they are supposed to know. So there is nothing I can really teach to the test.…It is very general, so you just have really good reading program and math program and they will be okay with a little test prep …throwing in there" (Interview on December 8, 2008).

In contrast, Mrs. Green and Mrs. Sommers at Central Elementary School were not implementing *Everyday Mathematics* as prescribed, and they heavily supplemented *Everyday Mathematics* with other materials. They found students were not ready for the activities required by *Everyday Mathematics*. Although Mrs. Green liked the approaches suggested by *Everyday Mathematics*, she often found her students were not equipped with basic skills and concepts that the students should have learned in previous grade levels. Mrs. Sommers found many of her 4th grade students could not perform multiplication, so she had to go over simple computation by using worksheets for practice rather than teaching *Everyday Mathematics.* Mrs. Sommers and Mrs. Green both mentioned that they were often stuck at Math Box, the review section at the beginning of the lesson. Both teachers used direct instruction to supplement these missed skills, which took a long time. Neither Mrs. Green nor Mrs. Sommers viewed *Everyday Mathematics* as an effective curriculum for their students.

This existing teaching practice of focusing on specific skills rather than following *Everyday Mathematics* seemed to have allowed Mrs. Green and Mrs. Sommers to incorporate the policy messages of teaching for tested standards. As DC BAS results returned, and as the DC CAS approached, Mrs. Green and Mrs. Sommers reported that they spent more time —focusing on standards that are fundamental" (Interview on

February 25, 2008), "attacking how to analyze questions (on the test)" to "gear our questions towards the test (sample questions)" (Interview on January 10, 2008) .

The staffing and scheduling seemed to have shaped the implementation of *Everyday Mathematics* at both schools. While Jefferson Elementary School had resources to support implementation of this reform-oriented math curriculum, teachers at Central Elementary School did not have a similar type of support. This difference seemed to have shaped not only whether teachers implemented *Everyday Mathematics* with fidelity, but also whether teachers viewed *Everyday Mathematics* as an effective curriculum for their students.

At Jefferson Elementary School, a Math Specialist provided intensive support for teachers to implement *Everyday Mathematics* by demonstrating lessons and providing suggestions to the teachers during the first two years of implementation. Both the Math Specialist and the teachers reported that the implementation was not smooth. The Math Specialist reported that he "begged teachers to put number lines in their classrooms." The teachers at Jefferson Elementary School reported that challenge was not only to learn new curriculum but also to respond to students. One teacher, Mrs. Reed, who closely followed Math Journal reported as follows:

> "The beginning part, there was a lot of… resistance with the kids, because their math had been different then. All the sudden, you have this and you are going to have to do multiplication in three different ways. They were not used to it. And I know the first year, a lot of kids, they hated math. Hated, really….time is going on…the third graders came to me …they are familiar with it, it is not so much resisted" (Interview on January 2, 2008).

When I conducted my interviews, during the third year of implementation of *Everyday Mathematics*, all teachers at Jefferson Elementary School mentioned that students liked *Everyday Mathematics,* and it worked well with their students.

At Central Elementary School, teachers did not have this type of on-going support. Hence, voluntary *Everyday Mathematics* professional development was the sole way for the teachers at Central Elementary School to receive additional training, which not many teachers regularly attended. Furthermore, teachers at Central Elementary School did not receive all materials needed to implement *Everyday Mathematics*. For example, the teachers did not receive white slates, which *Everyday Mathematics* recommends, while the teachers at Jefferson Elementary School had. Mrs. Sommers did not inherit the assessment books from a previous teacher, meaning she had to use her own worksheets or find worksheets from other curriculum.

Another factor that seemed to have shaped the difference in implementation of *Everyday Mathematics* at the two schools was the staffing and scheduling at these two schools. The instructional time for math was limited at Jefferson Elementary School, while the teachers had more latitude about how much time to spend on math at Central Elementary School. At Jefferson Elementary School, a math lesson lasted for 45 minutes each day for 3rd and 4th grade students. The teachers used this time for the main lesson of *Everyday Mathematics*. The teachers used a part of 30 minutes in the morning to do either math or reading skill building, often by using Math Box. This schedule was consistent throughout school year because there were special subjects scheduled, and students rotated to these subjects taught by different teachers. For example, for 3rd grade, Mrs. Caracelli did 30 minutes of mini-lesson of Math Box in the morning every

other week, and she taught the main part of *Everyday Mathematics* for 45 minutes in the afternoon. Mrs. Caracelli reported that this was the agreement made among grade level teachers. For 5th grade, Mrs. Canright had one hour to teach mathematics, and she included a short review session within this time by using a white slate. This structured schedule seemed to have made teachers continue with *Everyday Mathematics* throughout the year.

The teachers at Central Elementary School had more flexibility about time they spent on reading and math lessons. Both Mrs. Sommers and Mrs. Green had their students almost all day because special subjects did not happen on a daily basis. Students at Central Elementary School had only 45 minutes for science and music during each week, and 45 minutes of library time every other week. The school did not have a set time for social studies. It was largely up to classroom teachers if they wanted to provide another 45 minutes of science instruction to supplement science lessons provided by the Science Specialist. Thus, Mrs. Sommers and Mrs. Green could extend their math lesson if they wanted. This flexibility given to the teachers seemed to have contributed to the lack of implementation of *Everyday Mathematics* at Central Elementary School because Mrs. Sommers and Mrs. Green could supplement worksheets as they wanted, and there was not a clear agreement and monitoring within the school about how much time each teacher should spend on Math Box and on the main activity of *Everyday Mathematics*.

Summing up, a school‘s support system and scheduling seemed to have contributed to the different degree of implementation of *Everyday Mathematics*, an existing curriculum. This difference of implementation shaped the degree of instructional change among teachers. While teachers at Jefferson Elementary School

tended to continue with *Everyday Mathematics*, the teachers at Central Elementary School taught for tested standards.

9.3.2.4Experience of the Use of Assessment data

The difference between Central Elementary School and Jefferson Elementary School included existing school level practice using assessment data. While at Jefferson Elementary School, Math and Reading Specialists had worked with the teachers to identify instructional support for individual students by using the results from a different type of standardized assessments for years, teachers at Central Elementary School did not have this kind of experience. This difference seemed to have contributed to teachers‘ notions of the usefulness of the DC BAS results. The teachers at these two schools had different views about how much DC BAS should inform their instruction. Consequently, the teachers at Jefferson Elementary School and Central Elementary School changed classroom instruction differently in response to DC BAS results. While teachers at Jefferson Elementary School used the results of DC BAS almost like a checklist to make sure they covered the content with the understanding that DC BAS does not inform all aspect of students‘ academic performance, teachers at Central Elementary School thought DC BAS was helpful because it informed what to teach to make students perform well on the test. As a result, DC BAS guided instruction at Central Elementary School.

Jefferson Elementary School had administered their own assessment for a long time, and the teachers used this assessment as a primary resource to understand their

students‘ academic progress. For years the Math and Reading Specialists administered the assessment every four months, and they went over the results with classroom teachers, identified the students who might be lagging behind, and proposed supports the students might need. The Math and Reading Specialists provided extra support for the students by monitoring student progress, teaching students in small groups, providing individualized materials, and facilitating enrichment activities. At Jefferson Elementary School students‘ progress was monitored in relation to grade level expectation, which the assessment diagnosed. When DC BAS began, the Math and Reading Specialists went over DC BAS results with teachers, and they explained to the teachers that the benchmark tests could include topics students had not learned because DC BAS was to predict DC CAS performance at the end of school year. The school continued with the existing assessment system. Since the teachers had seen individual support provided to students based on the existing assessment system, the teachers did not see how DC BAS could help them understand where students were. The teachers also did not think that DC BAS provided them with information about how they should teach to improve student academic performance.

At Jefferson Elementary School, three out of four teachers, except for Mrs. Canright, continued with teaching *Everyday Mathematics.* Mrs. Canright did not follow the activities and the order prescribed in *Everyday Mathematics*; however, the degree of incorporation of the content suggested from DC BAS did not dominate her lesson. Mrs. Canright did not believe that DC BAS helped her to know individual students‘ progress, either. She reported that the existing assessment was much more helpful because she could see where individual students were in relation to grade level expectation.

Central Elementary School did not have this type of history of using assessment. Central Elementary School did not have a Math or Reading Specialist. At Central Elementary School, pulling out, individualized instruction, and enrichment activities did not happen as a regular part of schooling. DC CAS was the only test used to identify where students were in relation to the standards. Although Mrs. Sommers and Mrs. Green at Central Elementary School also noticed that the DC BAS results did not necessarily reflect the content they had covered in their regular instruction, the teachers extensively used DC BAS to guide their instruction because they heard DC BAS predicts how students will perform on DC CAS. The teachers thought that making students master tested standards was the way to help students perform well on the DC CAS. Thus, the teachers used DC BAS as the guide for deciding what to teach in their regular lessons. .

Summing up, a school‘s existing practice of using assessment to identify individualized support for students shapes teachers‘ views of the usefulness of DC BAS for informing their instruction. The teachers at Jefferson Elementary School did not see DC BAS as helpful for informing them about what support students should receive; thus, the teachers used DC BAS as the way to check off content to be tested. At Central Elementary School teachers viewed DC BAS as the guide of their instruction because they thought the information was helpful because it told them what students should learn. The teachers at Central Elementary School did not have exposure to different types of assessment that revealed or explained student progress.

9.3.3Teachers' Knowledge

As previous studies found (Coburn, 2002, 2004; Cohen, 1990; Spillane, 2000) a teacher‘s view of good instruction shapes a teacher‘s response to policy pressures associated with DC CAS. The teachers incorporated the policy messages about instruction that were in line with their pre-existing instructional approaches and beliefs about good instruction. Furthermore, as suggested by the Cognitive Approach, existing teachers‘ views of good instruction and their instructional practice shaped which policy messages they noticed and which policy messages they incorporated. In addition, the teachers‘ views of good instruction also shaped their views of how best to prepare students for the test. Although all teachers introduced the concepts and skills needed to do well on the test, when they did so, teachers kept pre-existing interaction patterns established in the classroom; thus, they all assimilated the content into their pre-existing instructional practices.

Teachers‘ existing views of good instruction shaped their understanding of how DC CAS and SAT 9, both high stakes test, differ. The teachers who taught close to traditional instruction, Mrs. Sommers at Central Elementary School and Mrs. Miller and Mrs. Reed at Jefferson Elementary School, noticed that the difference was in the degree of alignment of the content between the curriculum and the test. These teachers reported that their concern about preparing students for the test was that the test might include content that students did not come across; thus, they focused on covering the content and skills that might appear on the test. For example, Mrs. Reed noticed that DC BAS problems included multi-step problem solving, but she did not view it as a reflection of

instructional practices promoted by *Everyday Mathematics* and the Standards. *Everyday Mathematics* and the Standards promote students working on complex problems so that students could use multiple strategies. When Mrs. Reed went over the DC BAS problems with students, Mrs. Reed made sure that students knew each step of problem solving. The advice Mrs. Reed gave to students was to read problems carefully so that they would know if they had to do multi step problem solving. Mrs. Sommers viewed DC BAS problems as providing the format of word problems. Thus, she focused on teaching students how to find clues in the word problems that indicate which computation to use. None of them reported that they saw the nature of the problems on DC CAS as reflecting a different instructional approach.

The teachers who employed reform-oriented instruction, Mrs. Canright and Mrs. Caracelli, noticed that the nature of questions asked in DC CAS was different from SAT 9. Both Mrs. Canright and Mrs. Caracelli viewed DC CAS as close to reform-oriented math. Mrs. Canright reported that the questions on DC CAS ―gravitated to more real life situation…that is a valuable thing to teach" and ―it has been positive" for her instruction. Mrs. Caracelli believed that DC CAS was a good indicator of how students were doing on the fundamental basic 3rd grade level, and her students would do fine on the test with good curriculum, *Everyday Mathematics*.

The teachers' existing practices and beliefs about good instruction also influenced which policy messages they incorporated. The two teachers' contrasting responses to the policy message that teachers should encourage students to answer constructed response on DC CAS illustrates how teachers' views of how students learn math shaped whether the teacher incorporated this message into their actual instruction.

While Mrs. Reed, the teacher who believed students should master the procedure of problem solving, decided not to teach for constructed response, Mrs. Canright, a teacher who believed students should communicate math concepts, incorporated it into her regular class assessment.

During the school year 2007-08, the district provided mandatory professional development on constructed response questions in reading and strongly encouraged teachers to make sure that their students would answer these questions because the points for these questions were three times that of multiple choice questions, and even incomplete answers could earn partial credit. The district also provided professional development on constructed response questions in math, but it was voluntary training. Among teachers I interviewed, only Mrs. Canright and Mrs. Sommers had been to the district sponsored professional development on constructed response questions for math[27].

As the following quote indicates, Mrs. Canright believed constructed response questions were in line with her views of good math instruction. "The constructed response is more writing skills – writing about math, which is so important. Describing what you are thinking and being able to articulate your thoughts, either in writing or out loud" (Interview on March 14, 2008). Mrs. Canright taught constructed response in her class and made her in-class assignments include constructed responses. During the school year 2007-08, Mrs. Canright reported that she looked at Blue Print and found,

[27] For Mrs. Green at Central Elementary School, she told me that she did not receive the math portion of the training. For teachers at Jefferson Elementary School, three out of four teachers I interviewed participated in the training by the Math Specialist in which teachers went over the students' answers and exchanged their understanding of how to interpret the scoring rubric.

—there will be constructed responses about measurement, within the area of triangle, so these things will have constructed responses. I have given my children written constructed response to prepare them for [the test]." (Interview on April 2, 2008)

Although Mrs. Sommers was interested in learning how she could teach for the constructed response and went to the training, she decided not to teach for the constructed response as DC CAS required. Rather, Mrs. Sommers focused on the multiple-choice questions. The following account presents how her belief of how students should learn math influenced this decision.

> —Because most of the tests are the objective kind…And, anyway, they [students] will not be able to do the (constructed) response test if they don't know the skills. So I have to do something about the skills, first. Because…they cannot explain it unless they know how to do it." (Interview on April 3, 2008).

Although Mrs. Sommers knew that constructed response could get a partial score and she totally agreed that she should make students perform well, Mrs. Sommers's view of how students should learn math shaped her priority for test preparation. She dropped the preparation for constructed response. The rest of the four teachers covered the constructed response; however, they did not report that they were spending more time on constructed response than in the previous year.

Teachers' view of good instruction also shaped their view of how best to prepare students for the test. Teachers who employed traditional instruction tended to think that students should be prepared for the test through repeated practice in multiple-choice test taking, through identifying mathematical computation from the language used in the problems, and through shortcuts. These teachers reported that they were concerned that

students may see problems, concepts, or terms that their lesson did not cover. Mrs. Sommers believed that running out of time was the reason some students would not score high. Hence, she taught students to use shortcuts. For example, Mrs. Sommers discouraged students to calculate by using expanded notation when they took the test although she taught the extended notation to make sure students understood the concept of place value. –I do not really take (Note: time for) kids to do expanded notation. Just...the important thing...just let them understand how it should be done. But then, I let them realize that it is too much time. If it takes too much time to do expanded notation…But there is a shorter way. Just teach them short (Note: shorter way) for test taking skills, because we are under time restrictions" (Interview on January 11, 2008). In the lesson, she taught students to just count the number of zeros rather than writing down expanded notation at the testing situation.

Teachers who employed reform-oriented instruction did not think that a specific technique could make students perform well on the test beyond providing a limited time of review of concepts students learned during the year and reminding them of basic test taking skills such as reading a question until the end of it. For example, Mrs. Canright did test preparation when DC CAS approached by making students create multiple choice problems and assigning students review questions as homework. Mrs. Canright went over concepts and skills that most students had trouble with. Mrs. Caracelli also did test preparation in which students reviewed math problems one week before DC CAS. These teachers did not see that there was any specific ways to prepare students for the test, and teaching shortcuts did not make sense to those teachers because students would not learn math in this way.

9.4 SUMMARY

Analysis of the six teachers found that all teachers knew that they had to improve students‘ performance on DC CAS, but high stakes testing did not change their instructional strategies. All teachers noticed the content of instruction from the policy pressures, and they taught this content in their classrooms. However, the teachers differed in their understanding of whether the problems on DC CAS reflected their instructional approach. The teachers who taught in a reform-oriented approach noticed that the problems on DC CAS were in line with their instructional approach. Other teachers thought that DC CAS asked them to cover specific content. Thus, the high stakes testing did not change the teachers‘ instructional approach, but it made teachers feel that they had to make students perform well on the test. Drawing upon existing studies from Cognitive Approach to Instruction, the study examined how the following dimensions shape the teachers‘ responses: the frequency of encountering policy messages, the strengths of policy messages teachers viewed, and the degree of concreteness of the messages. The teachers tended to notice messages that were concrete, the messages they came across frequently, and the messages they saw as strong messages.

The study also identified that a school‘s organizational context shaped teachers‘ responses. The school‘s improvement status enabled Jefferson Elementary School to be

buffered from the messages. In my study, principals, counselors, and a Math Specialist mediated the policy messages. Staffing, scheduling, and the existence of instructional support within a school shaped existing classroom instruction and teachers‘ views of how to use DC BAS results.

Analysis of policy pressures teachers had come across indicated that although the teachers received the policy message that they had to improve students‘ performance on DC CAS, the policy pressures rarely carried any message about how teachers should teach using the reform-oriented instruction that the standards and *Everyday Mathematics* promoted. The messages about the change of instructional approach was carried at the individual level through conversation or training, and there was no monitoring of actual instructional practices in the classroom, while the number of students who needed to be moved to make AYP had been periodically monitored. The next section presents how science low stakes testing influenced teachers‘ classroom instruction.

10.0 FINDINGS: CHANGE OF SCIENCE INSTRUCTION

In this section, I describe how science teachers changed their classroom instruction in response to science testing. The low stakes science test was introduced for 5th graders during the school year 2007-2008, which was the year I collected data for this study. I collected data from three teachers, Mr. Berry, Mr. Shawver, and Mrs. Green (who were responsible for teaching 5th grade science) to understand how science DC CAS influenced their instruction. Mr. Berry and Mr. Shawver were science teachers who had been teaching science for multiple years. Mrs. Green was a classroom teacher, and she began to teach science in 2007-08. I interviewed Mrs. Green to find out about her instructional change in math and science. Mr. Berry and Mrs. Green taught at Central Elementary School, and Mr. Shawver taught at Jefferson Elementary School.

The outline of this chapter is as follows. I first present how these three teachers' instruction changed over time. Policy pressures associated with the science test were almost non-existent, and only one teacher, Mr. Shawver, reported that he made any change in instruction. Mr. Berry and Mrs. Green at Central Elementary School reported that their science instruction changed in response to math and reading testing. Second, I present how policy signals, context, and teacher knowledge shaped responses to policy pressures associated with DC CAS.

10.1 THREE TEACHERS' CHANGE OF INSTRUCTION

In this section, I present each teacher's view of good science instruction, their encounters with science standards and science DC CAS. Three teachers reported that the pressure to improve student scores on the science test was almost none, and they did not think that their methods of instruction were influenced by DC CAS. However, their views of science instruction were different, and they responded to different policy pressures. While all teachers believed in the importance of using real life materials in science lessons and in the importance of promoting students' interest in science, the degree to which they taught scientific concepts and techniques varied. In addition, the degree to which science lessons were buffered from pressures to teach for math and reading tests varied. At school Central Elementary School, science teachers began to teach for math and reading tests, while Mr. Shawver at Jefferson Elementary School could continue teaching science.

10.1.1Mr. Shawver at Jefferson Elementary School

Among the three teachers, only Mr. Shawver incorporated policy pressures associated with science testing. He spent some of his lesson time going over problems similar to those he saw on the pilot test with students. As with the other two science teachers, he received only a small number of policy pressures associated with science DC CAS.

Mr. Shawver began teaching in the school year 2004-05 at Jefferson Elementary School. The year I collected the data, school year 2007-08, was Mr. Shawver's fourth year teaching at Jefferson Elementary School. He had been teaching the 5th grade since he began teaching at Jefferson Elementary School. Jefferson Elementary School took a team teaching assignment for the 5th grade. Mr. Shawver taught science and social studies for all 5th graders for one hour every day. He either split the one-hour period into science and social studies, or he used one hour for either subject depending on the activities students did. Students typically spent one hour for science if they did experiments or observations.

Mr. Shawver's view of good science instruction is close to the instruction promoted by DC Standards. Mr. Shawver believed science should be "hands on," and students "should be interacting with the materials and the world around them, " and he believed these activities would promote interest in science. Mr. Shawver also reported that students' interest drove his instruction because he believes that if students are motivated, then they learn better and classroom management won't be a problem. Mr. Shawver said students should become skilled at "asking good questions, and trying to find answers to those questions" in science class. As the following account indicates, "asking good questions" entails students being curious about their surroundings so that this curiosity could lead to scientific investigation.

> "We are doing the sun and the moon recently, and why do we have the day and night? Why do we have seasons? A lot of times there is stuff they have seen out there, if you put it to them, they start thinking about it and make sense of it, make

connections between what they see and what they know and try to form an idea or the theory about why things are the way they are" (Interview on March 14, 2008).

Mr. Shawver also believed that students should conduct science experiments and observations. For example, Mr. Shawver reported that a science fair that all 5th graders at Jefferson Elementary School participated in provided this opportunity because students ~~–m~~ade their own science experiments from hypotheses" (080314). The following account indicates that Mr. Shawver asks students to learn scientific investigation.

> ~~–I~~ should ask them, in some way, how dependable is that answer? So you have some ideas about how the data is reliable, if you make decisions based on the reliability of the data, how good is the decision…they should ask questions that are, like, well in an experiment, if you are trying to test a hypothesis, you have to form a question in a certain way, or make sure the question and the experiment are connected, so that you can answer the question with the experiment you are doing" (Interview March 14, 2008).

Mr. Shawver planned his instruction around the FOSS kit that had been used in Jefferson Elementary School when he came. He thought the FOSS kit was a good curriculum because students learn science through hands on activities. In his classroom, which was at the corner of the open space, four students shared a round table. Students often worked in groups by doing an observation or experiment and coming up with conclusions as a group.

Mr. Shawver assigned long-term projects from time to time. One project asked students to record the time of sunrise and sunset every Monday. Each student recorded the information on a one-page paper by using red and blue bars, which showed the

change in the lengths of night and day over time. The following account illustrates several purposes of this project.

> –They need to go and find that information, internet or newspaper, so that is one piece. And they have to record it…. Graphing the information is another piece. And they can see the pattern of the days getting shorter and then days getting longer, you connect that with seasons changing….I just feel that is satisfying because it is simple and they are just coloring, they just use two colors, but you are able to see it. Days getting shorter and longer and compare how long it is. It is a graphing skills and data collection. Something they are responsible for" (Interview on March 6, 2008).

Mr. Shawver thought that students with lower level reading skills had a more difficult time. For those students, he tried to incorporate hands-on activities as much as possible to make students interested in scientific concepts. The following accounts indicate that Mr. Shawver believes that even students who do not have high level reading skills can learn science.

> –Some of the kids are really good at reading the textbook, and they can learn information from the textbook. Earlier in the year we were doing more hands-on stuff. I think most of the kids like that. So I have been trying to do that more…Some kids are good at science but cannot read very well, so the reading part will frustrate them. So, that's another reason why I like to do hands-on labs when I can, and I make the time to put them together." (Interview on March 6, 2008)

Although Mr. Shawver was familiar with the overall concept of the new standards, he did not think that the standards influenced his science instruction. Mr. Shawver received training on the standards when math and reading standards were introduced in the school year 2005-06, but he did not attend training for the science standards in 2006-07 because he attended the training for social studies instead. In these professional development sessions, he learned how to identify power standards and how to find vertical alignment. Mr. Shawver reported that that this professional development did not change how he planned and taught science and social studies lessons. Mr. Shawver reported that he continued with teaching the FOSS kit. When the new standards came in, he looked at the standards and report cards to decide what content he needed to cover. If the FOSS kit did not have the content written in the standards, he used other materials such as other books, or videos. In 2007-2008, Mr. Shawver said he rarely used the new textbook. Mr. Shawver‘s report indicates that the standards had influenced the content of science instruction.

When asked about the influence of the standards, Mr. Shawver reported that the professional development activities might have made him use a specific approach to communicate with other subject teachers about what students are learning in science class. In the professional development, a group of teachers listed concepts and skills students learn from earlier grades to older grades to identify when what types of concepts and skills should be learned. Mr. Shawver reported that he used a list of concepts and skills from the science standards to communicate with art and music teachers what his students were learning in science class. He reported that this way of communicating

students‘ learning with other teachers by using common language might have come from the professional development activities. At the same time, Mr. Shawver also added that team teaching at Jefferson Elementary School encouraged him to coordinate instruction with teachers of other academic subjects, particularly reading and math. Within the team-teaching situation, Mr. Shawver discussed how the role of science and social studies is to reinforce concepts and skills students learn in math and reading because ―they make practical use of math and reading, and make the concepts learned in reading and math less abstract, more concrete" (Interview on March 14, 2008). He reported that he pays more attention to what is happening in reading and math lessons so that he can use the same vocabulary to reinforce the concepts and skills.

Mr. Shawver also reported that the impact of the science test on his instruction was minimal. Fifth graders took a pilot science DC CAS in the spring of 2006 (the end of school year 2006-2007). Before the pilot test, he received a booklet, which he thought was the collection of Massachusetts‘ assessment. Mr. Shawver went over problems with the students. Mr. Shawver did not receive the results from the pilot test, so he did not know how his students did. In the school year 2007-08, Mr. Shawver received Blue Print. He reviewed it to check what would be tested on the DC CAS, ―but not like with a magnifying glass" (Interview on January 9, 2008).

Since the pilot test, he said he was paying more attention to the test and spending a little more time on the topics that appeared on the test. He said the following:

—Maybe just mentioning or making more of a point of it, if something like that comes up when I am teaching, I spend a little bit more time on it. Like graphs, or something comes-up, maybe I slow [the lesson] down a little bit. I do not think I made a huge adjustment based on the test, but just I am trying to be more aware of it" (Interview on January 9, 2009).

Mr. Shawver saw there were problems about planets on the pilot test last year, and the following account indicates that Mr. Shawver was paying attention to the content that appeared on the test.

—I am trying to put a little more focus on vocabulary. I updated the science vocabulary words, have them up and try to refer to them --- just on concrete things that they can memorize….I imagine in their brain they connect to the things that they might know and see a lot. Also, I just remember the standardized test last year had a lot of astronomy. A lot more than anything else…no, I did not take notes but what I remember was there was a lot of astronomy" (Interview on March 6, 2008).

He remembered it because his students did not learn astronomy until the end of the school year, but it showed up on the pilot test. When he taught the topic of astronomy, he made students come up with mnemonic devices to help them memorize the names of the planets. He believed this activity was for memorization, but at the same time he thought this activity also provided students with strategies for test-taking. He assimilated DC CAS into his instruction by using approaches he already implemented in science. He did not change his instructional approach.

Mr. Shawver reported that although there are some basic skills students might need to know, such as reading problems carefully or going over all answer choices, he did not believe that the standards test reflected students‘ ability, and he said he is careful not to teach too much ―for the test.”

> ―I do not believe the focus on testing makes better children, better humans when they are grown up...so I am always trying as little as I can. I think test-taking strategies are okay. I mean there are some things that overlap with other things we are doing, slowing down, re-reading their work, going over the answer, what kind of things are more like study habit? But, I do not put too into picking apart questions” (Interview on March 6, 2008).

10.1.2 Mr. Berry at Central Elementary School

Mr. Berry did not change his approach to science in response to policy pressures associated with science testing. However, Mr. Berry introduced reading and math test preparation in his science lessons. The below description of Mr. Berry‘s instructional change shows how the organizational context of schools, scheduling, and the policy pressures associated with high stakes testing influence science lessons. As with Mr. Shawver, Mr. Berry continued with the instructional methods that he learned when he entered into teaching. Mr. Berry had a similar view of good science instruction as Mr. Shawver: that students should learn science in relation to their everyday lives and that science class should get students interested in learning about their environment.

However, his instruction did not cover activities to make students learn scientific investigation as Mr. Shawver did.

Mr. Berry started teaching in the early 1990s after he had worked as a lab technician for 15 years. He taught, until two years ago, at another elementary school that had a strong emphasis on science and math education. It was the second year for Mr. Berry at Central Elementary School when I interviewed him. He had been a Science Specialist Teacher throughout his teaching career. During the school year 2006-2007, his first year at Central Elementary School, he taught older grades twice each week, but in the school year 2007-2008 he was teaching science once per week from kindergarteners to 6th grade. This schedule change derived from a decision the school leadership made to keep older students with their homeroom teachers in the morning.

Mr. Berry planned his science lessons around what he viewed as important scientific concepts. He drew upon the curriculum that was in place when he began teaching in 1992, i.e., "Competency Based Curriculum" which described skills and concepts students should master as the framework to plan lessons. He viewed the CBC's description of learning objectives as being specific and easy to understand for students. . "It is good, I mean it is objective. The children know what they are trying to learn with CBC. That is very simple as opposed to the standards (Note: current standards) that were very long." (Interview on April 4, 2008). However, he did not follow all contents in the Competency Based Curriculum. Rather, he selected concepts that he viewed as foundational for students to continue to learn science at the middle school level and beyond.

Both Mr. Berry and Mr. Shawver believed that science instruction should promote students‘ interest in science. As the following account indicates, Mr. Berry thought that science should be taught in relation to everyday life and that the most important thing in a science lesson is making students curious about their environment; these values shaped his decision about what to teach in a lesson.

> –I teach science with the multi-disciplinary approach, trying to blend in skills and knowledge of students – In other words, information that they already know and show them that they know more science from their world than they think they know. I want them to understand that science is not something that you just do at two o‘clock on Tuesday. It is everyday. It is an integral part of your life. It is very important…It is not difficult. It is fun -- and it is something you already know. It is my job to bring it all together for them, these individual pieces” (Interview on December 7, 2007).

As with Mr. Shawver, he reported students‘ interest drove his lessons. Mr. Berry and Mr. Shawver both fielded students‘ questions during the lessons so that students could see that their questions about everyday life were discussed in the classroom.

However, Mr. Berry differed from Mr. Shawver in his views of the goal of science instruction. Mr. Berry viewed the goal of science instruction as providing a good foundation by making students notice that there is a connection between their life experience and scientific concepts, so that students would continue to be interested in science. However, as the following account suggests, contrary to Mr. Shawver, Mr. Berry does not believe that students need to master the procedure of scientific investigation. For his science lesson, he says the following:

—[students] do not have to master [concepts or vocabulary] because my focus here is to make them start thinking scientifically and to see science in their lives everywhere. So, when they get home tonight and their mother is cooking in the kitchen and they see the steam, _oh, you know we did something like that at school, that is because it is hot and 100 degrees.' That is the level. I do not try to do so much with all this vocabulary and spelling, paragraphs, structure and all that. I just want the kids, by the time they go to junior high school, to have some ability to reason scientifically" (Interview on April 18, 2008).

In contrast to Mr. Shawver's students who designed a science experiment, students in Mr. Berry's class received only an introduction to scientific investigation. The range of activities students did during Mr. Berry's class was talking about what they know about specific scientific phenomena and taking notes on experiments or observations Mr. Berry conducted experiments in front of students. Mr. Berry rarely assigned a long-term project, and only a couple of students participated in the science fair. Mr. Berry's students did fewer activities that included creating hypotheses, conducting experiments, or enacting systematic observation. Students' activities remained at the level of observing the teacher's demonstration of an experiment or using a cylinder to measure a quantity of water by following the teacher's specific instruction. Mr. Berry's students get a glimpse or a snapshot of experiment or observation, but they do not conduct a whole experimental process.

Mr. Berry believes that students should take notes to document their observations and that note-taking is the way to learn scientific concepts. Mr. Berry learned note-taking

from his mentor teacher when he began teaching. In the following account, he suggests that:

> –She (Mr. Berry's mentor teacher) taught [students] about writing objectives and made them keep notebooks. She even had the kindergarteners doing that. She said it was good documentation for parents to look at. Probably kids will look at it. When they have the notebook they tend to own it. They can feel pride in it" (Interview on December 7, 2007).

Mr. Berry had taken this approach, making even kindergarteners keep notes in a composition notebook, which he also viewed as the approach to learn science because it made students carefully observe and document their observations.

Mr. Berry reported the challenge he had with his students was that they do not like to write down their explanations and observations. He reported that his students had a difficult time writing and explaining observations. He thinks that being able to explain their observations is an important skill students should have. To develop this skill in his students, he used note taking. –They are not really very academic in their approach to learning. If you let them draw something, if you allow them to talk about it, it is fine. (But) they do not like…explaining. I say yes, you need two paragraphs to explain that. I say you can always use pictures in your explanation…but you cannot have pictures that are unlabeled, so they got and do that...I kind of encourage them to explain it to somebody else…[students] got to be able to explain it. That is very important for me" (Interview on December 7, 2007). When I observed, about half of the instructional time was spent on students taking notes.

While Mr. Berry drew upon CBC and planned his instruction around scientific concepts that he thought were important, new science standards did not influence his instruction. As with Mr. Shawver, Mr. Berry also believed that science instruction was not standardized in the district. Mr. Berry described how he decided what to teach in the following way:

> "I would probably think that each science teacher teaches differently according to whom they have. We have standards, but we do not have a particular curriculum that says do this" (Interview on December 7, 2007). "[The district] does not tell us how to do it, they say, like book, standards, here is report card, just do it" (Mr. Berry 071207).

Mr. Berry also mentioned that standards, report card, and textbook are not aligned.

> "The report card is not aligned with the standards, and it is not aligned with the textbook. So, if you look at the report cards for a certain period –so did we touch this area? And this will generalize (Note: the activities students did in the classroom). Then you can teach. But to the degree of mastery or beginning of developing (Note: as mentioned in the new standards), things like that, nobody explained that to me" (Interview on April 4, 2008).

Rather than following science standards, Mr. Berry teaches the same topics across grade level at Central Elementary School. For example, I observed the second and third graders learning how to measure water by using different types of cylinders. The activities given to each grade level are slightly different, such as the 2nd grader used only one kind of

cylinder, while the 3rd graders used two different sized cylinders. The following account indicates how Mr. Berry taught the same concept to different grade level students.

> "I can get the concept in there --- kindergarteners can learn much the same thing as 6th graders. It is just not as complex of form for them, but it essentially is the same experience. And kids do not even realize it. That is how I present it to them…What I try to do is, I teach major topics, like solid, liquid, and gas, which is a major, a big theme they will talk about that even when they are in college. So I teach that and try to teach that really well" (Interview on December 7, 2007).

Mr. Berry's view of what counts as an important concept for science did not necessarily align with the standards. In the new standards, while 2nd grade physical science focuses on different states of materials, other grade levels introduce energy. Hence, Mr. Berry's view of how scientific concepts should be ordered across grade levels was different from the standards.

Science of DC CAS did not change Mr. Berry's instruction at all, but Mr. Berry incorporated test preparation for math and reading DC CAS. When the pilot science test was administered in the spring of 2007, Mr. Berry did not get any information about it, but he saw a sample test a couple of days before the field test. He noticed a difference between the assessment the district had in the early 1990s and the DC CAS pilot test. He reported that DC CAS requires math and reading ability to answer science questions. He explained as follows:

—(Note: Test in the early 1990s was) based on what children can do. Showing children pictures of things and want to know what comes next, what did you have to add to this make this? Now, the test is more focused on reading, and charts and graphs, and you have to be able to analyze and synthesize from what they read" (Interview on December 7, 2007).

Mr. Berry reported that he allocated some time from his science lesson for math and reading test preparation. The following account indicates that Mr. Berry believes that math and reading are important foundational skills for learning science.

—If they are having trouble with math, they might have trouble with some of the charts, and graphs, or measurement. It ends up the same thing. If they have a problem in any area of that test, it will show in science. You cannot separate these things….in many instances, if you have trouble with math and then we get into certain things....it is still quantitative reasoning. If you have trouble with measurement in math, there is measurement (showing a science test problem). It is the same thing." (Interview on April 4, 2008).

This year, he said, having DC BAS results helped him understand where students had trouble, and Mr. Berry reported that the DC BAS results were helpful for him and for schools in order to target specific skills in reading and math. Mr. Berry's account suggests that he came across policy pressures associated with high takes testing. Mr. Berry heard that the students at Central Elementary School had difficulties with comprehension of informational texts.

—[The counselor (Note: at Central Elementary School counselor presented test data to the teachers)] showed the trend -- weakness in (Note: comprehension of)

informational texts. If they have a problem with informational texts, they might have a problem when it comes to reading test problems. Yes, that was a weakness throughout this building. There was not much emphasis on that [last year], we did not necessarily know (Note: where students had weaknesses). We did not pencil in the specific weakness, but we have it down pretty much now ---that's informational text, which is about 75% of time (Note: students had trouble with)"(Interview on April 4, 2008).

Although students do not like going over the test preparation, Mr. Berry did the test preparation for reading and math in his science classes, and he viewed preparation as being more focused than in previous years. He allocated about 10 minutes for students to work on reading worksheets. The content of the reading did not necessarily relate to the topic students were studying on that particular day. He reported that as long as he sees some –science" in the math and reading problems, then he uses those problems. What Mr. Berry means by –science" seems to be any topics that may be mentioned in science lessons rather than making students learn any scientific procedure.

–We do a lot of workbooks, we do graphing, we do charts. I even had them identify operations. One time they said, _why are we doing this? This is not a math class.' I said, _I want you to get used to the fact that there is math in science. Sometimes we have to figure out temperature differences, 90 degrees and it is 30 degrees. How much difference is that? You have got to find that out.' So, me having them understand that it's all connected to reading to science to social studies, it is all connected" (Interview on April 18, 2008).

Mr. Berry did not receive information about what the science DC CAS would look like, and he did not prepare students for the science test. He sometimes looked at science tests from other states, such as Massachusetts, Maryland, and New Jersey to see ―the level of questioning, process of elimination, and whether it is constructed response," but he did not do test preparation for science except for going over one or two problems from the other states' assessments. Mr. Berry continued with math and reading test preparation. Mr. Berry changed his lesson in response to the math and reading tests, but not in response to the science test.

10.1.3 Mrs. Green at Central Elementary School

While both Mr. Shawver and Mr. Berry share similar views of good science instruction and they provided students with hands-on activities, Mrs. Green structured her science instruction around teaching reading skills. Although Mrs. Green used the new science textbook, the tasks she assigned to the students were intended to improve reading skills. She did not change her science instruction in response to the science test. Rather she continued to teach ―science lessons" that are actually reading lessons. Below, Mrs. Green's case shows how weak institutional pressures associated with science made Mrs. Green continue with reading instruction.

It was Mrs. Green's first year teaching science when I interviewed her during the school year of 2007-2008. During previous years, she did not teach science because Mr. Berry, a Science Specialist Teacher, taught her students twice each week. Before Mr. Berry, there was another Science Specialist Teacher at Central Elementary School. She

had done some science-related activities when the topics of the reading textbook covered science related topics, but she had not taught science –straight" (Interview on December 11, 2007). But this year, students had science only once per week, so Mrs. Green began to teach science by using the new science textbook that the district adopted this year (2007-2008 school year).

Mrs. Green reported that she did not feel that there was a pressure to teach science although she believes science is important. For example, Mrs. Green reported that she still did not know how much time teachers should spend on teaching science. In addition, Mrs. Green reported that because there was a pressure to perform well on the math and reading tests, she did not feel she could spend time on science. She reported that she often integrated science content into reading and math. She said she mostly teaches science by integrating it into reading, which is also recommended by the science textbook.[28] The following account indicates that integration for Mrs. Green means covering science content by teaching reading skills.

> –I normally integrate science into something else. Like today I taught science, but I integrated into writing so we were writing. We learned about volcanoes and they write, science, a big paragraph with main ideas and supporting details, which are facts." (Interview on December 4, 2007).

When I observed her science lesson, students made a study guide, which was a summary of each section of the textbook. As students worked individually, she told

[28] —A Closer Look" s teacher's manual provides suggestions for teachers who have limited time for science. The manual presents two options of how teachers can allocate time. The first option includes doing the experiments explained in the book, and the second option does not include the experiments. The students' activities suggested in the textbook are reading the chapters. Hence, its approach to science instruction is different from FOSS, which recommends a hands-on approach to science education.

students that they should write down vocabulary and definitions, as well as think about what the main ideas in each chapter were; students took the study guide home to review[29]. She explained that the activities they did for the chapter they were on as follows:

> "I want to make them understand the main idea…I also have them do the main idea and supporting idea wheel. An idea wheel has the main idea in the middle and the details around the side. So they get an idea of what that particular chapter was all about" (Interview on March 13, 2008).

For other science lessons, she planned lessons around "compare and contrast."

> "We compared and contrasted the energy that hits the earth. What was absorbed and what was reflected. We compared and contrasted global winds to local winds. But before we do that, I had them identify the main idea and details. Because it is going to be very difficult for you to compare and contrast something if you don't understand, I have them write down the main points. So I have them always do the main idea in detail before we do anything else" (Interview on March 13, 2008).

Mrs. Green reported that from her experience students understand scientific concepts better when "they can first identify main ideas clearly" (Interview on March 17, 2008), suggesting that she viewed reading as the foundation for learning science. Her view of how students should learn science did not seem to be established as much as her view of how students learn mathematics. Although she reported that she thought students

[29] At Central Elementary School, students do not bring home textbooks because they might lose them.

—should learn scientific process and students should be able to apply scientific process to their life," (Interview on March 17, 2008) her science instruction was structured around reading skills.

Although she sometimes made her students work in small groups to do science experiments that appeared in the science textbook, she did not facilitate discussion about their observations to connect them to scientific concepts.

Mrs. Green received a minimal amount of policy pressures from the Science Standards and Test. Mrs. Green remembered that students took a pilot test last year. —It did not mean much. They just did it just to see how it went" (Interview on December 4, 2007). As the following account suggests, there is not much discussion about science instruction at the school and district level, and Mrs. Green did not feel pressure to teach science.

> —We have a curriculum and so, and let me state this, at times I do teach straight science. It does not happen as much as I'd like, but there is definitely a time I do science from the science textbook. So, I mean there are those times, but they are rare, like once a week" (Interview on December 11, 2007).

Mrs. Green did not remember getting training about science standards besides receiving a laminated poster-sized chart of science standards for 5th grade. When she taught science in 2007-08, she did not refer to the standards. —Honestly I do not know the science standards. It is unfortunate. What I do (is), I focus on reading standards. I think there is a pacing guide in the textbook. I think there is, but (laugh) I definitely do not follow it" (Interview on December 11, 2007). She liked this new science textbook

because it describes how to integrate reading and math standards into science, and it comes with charts and transparencies to be used in lessons, which saved her time. Although the textbook also suggests experiments, she says it is difficult for her to do the experiment suggested because they do not have time, and she does not have all the suggested materials. She wishes she could. Mrs. Green's approach to science teaching was different from what the district intended. The district intended the science teachers to design instruction around the standards; however, Mrs. Green did not have opportunities to learn the standards.

Mrs. Green hoped that she could collaborate more with the Science Specialist Teacher, but she found it difficult. When she talked about collaborating with the Science teacher, she meant to be dividing instruction between vocabulary and science, which also indicated that she viewed reading strategies as foundational for science learning. She did not hear anything about science DC CAS, except for receiving Blue Print, and she did not look at it. Rather, she continued focusing on teaching reading skills.

After the DC CAS ended in late April, 2008, she began a project in which students learned about environmental issues. She said this project included reading and science portions. ~~–~~It would be total integration. The purpose of the project would be to raise students' awareness of global warming" (Interview on April 7, 2008). During the month-long project, Mrs. Green introduced global warming, and students read about climate change, and learned how they might change their life habits to be more environmentally friendly. She prepared worksheets, showed power points, and video clips, to help students understand various issues associated with global warming. At the end of the project, students brainstormed what they could do for conserving energy, and

Mrs. Green made Power Point slides based on the students‘ ideas. Although the topic covered issues around environmental education, a review of the fieldnotes suggests that activities students did were focused on extracting information, rather than developing scientific concepts. There were no activities for designing and conducting experiments or observations. Students‘ activities centered around listening and summarizing information provided by Mrs. Green, which indicates that her ideal science instruction is also shaped by her reading instruction.

The case of Mrs. Green shows that her science instruction did not change in response to policy pressures associated with the test because the pressures to teach for the science test were almost non-existent. As with Mr. Berry, Mrs. Green incorporated reading into her science lessons.

10.2ANALYSIS

All teachers reported that they did not see science testing as changing their instruction. All teachers came across very few policy pressures associated with the science test (i.e. Blue Print for the three teachers and a seeing Pilot Test for Mr. Shawver and Mr. Berry), but the three teachers were different in degree of incorporation. Mr. Shawver spent some time on the topics he found on the pilot test. Mr. Berry did not change science instruction at all, but he allocated some time for doing reading and math test preparation. Mrs. Green did not do any test preparation for science, but she taught science content

around the reading and math skills that she thought were important for doing well on math and reading DC CAS.

Policy pressures associated with DC CAS Science did not change the teachers‘ existing instructional practice. The instructional approach Mr. Shawver and Mr. Berry used for test preparation was in line with their existing instructional practice, although both of them were aware that the content they were teaching was separate from their regular lessons. Mr. Berry went over worksheets with students, and Mr. Shawver made students create mnemonic devices. Mrs. Green taught reading skills, which was continuation of her reading instruction. For all three teachers, their existing science instruction did not change as they came across DC CAS.

In this section, I discuss how policy signals, contexts, and teachers‘ beliefs shaped the way they incorporated policy pressures associated with science DC CAS.

10.2.1 Policy Signal

The teachers did not come across policy pressures associated with the science test that carried messages that their students must pass certain science-related performance measures. Mr. Shawver and Mr. Berry saw the problems of the pilot test in spring 2007, but they did not see the results of the pilot test. Mrs. Green did not come across any policy pressures associated with the science test, except for occasionally seeing the sample test from other states on-line. The district did not administer the benchmark test, DC BAS. There was no improvement target for science. The teachers did not experience the pressure to improve performance on the science test. This weak and very infrequent

institutional pressure seemed to have shaped the teachers‘ responses of ignoring (for Mrs. Green and Mr. Berry) and spending short time on going over test problems (for Mr. Shawver).

Table 10 List of Policy Pressures and Teachers' Response

Policy Pressures Associated with:	Teacher	Policy Pressure	Carrier	Response	Aspect of instruction influenced by the policy pressure	Response (in relation to existing instructional practice)
DC CAS Field Test	Mr. Shawver	He saw the field test when he administered	Field test	He remembers there were many questions on planets. He introduced vocabularies in his lesson by making students create mnemonic devices.	**Content**	**Assimilation**
2008 DC CAS Science Test	Mr. Shawver	He received Blue Print.	Blue Print	Looked at it, but not paying too much attention	No specific impact on instruction	**Rejection**
DC CAS Field Test	Mr. Berry	He saw the field test	Field test	Saw it, and thinks that it was different from Competency Based Curriculum Test. He saw graphs, and thinks math is integrated. He also checks to see what types of problems and what depth of knowledge are asked	No Impact to Science Lesson, Introduced 10 minutes math and reading worksheet	**Rejection** **Parallel Structure**
2006-07 Science DC CAS Pilot test	Mrs. Green	Administered DC CAS Science Pilot Test	Field test	—It did not mean much"	No impact on her science class	**Rejection**

10.2.2 Organizational Context

School‘s improvement status, staffing, and the availability of knowledge resources in the school seemed to have shaped the differing responses among teachers. While Mr. Shawver at Jefferson Elementary School did not allocate his instructional time to do reading and math test preparation, Mrs. Green and Mr. Berry at Central Elementary School incorporated policy pressures associated with the math and reading tests. Whether the teacher incorporated math and reading pressure into the science lessons seemed to have derived from the following difference between school Central Elementary School and Jefferson Elementary School.

The AYP status of these schools shaped what policy pressures teachers encountered. As I described in the previous chapter, Jefferson Elementary School was buffered from policy pressures that carry the message that teachers should teach tested standards, because teachers at Jefferson Elementary School were exempted from the professional development by the district, and the principal and Math Specialists facilitated the activities that did not carry the message that they should teach for tested standards. The teachers at Jefferson Elementary School did not come across policy pressures that require a specific number of students to improve in math and reading.

As Central Elementary School was in improvement status, the teachers received more policy pressures indicating that they had to move up a certain number of students to the proficient level, and the school leaders presented DC BAS results as the way to guide instruction. All teachers at Central Elementary School were involved in providing math and reading test preparation. Mr. Berry and Mrs. Green saw DC BAS results and analysis, and Mr. Berry

thought the test preparation had become more meaningful and focused this year. Mr. Berry and Mrs. Green were exposed to strong policy pressures to improve math and reading.

Staffing and scheduling seemed to have made science lessons continue at Jefferson Elementary School. At Jefferson Elementary School, science and social studies classes were scheduled as separate academic subjects from reading and math, and science class was viewed as an important subject area to provide students with hands-on activities. The students rotated from math class to science class, and all students attended a school-wide science fair. In addition, because the 5th grade teachers shared their students, teachers communicated frequently about what students were learning, and art and music teachers often approached classroom teachers to learn what concepts, skills, and topics students were learning in other academic subjects so that they could incorporate them into their lessons. This frequent communication seemed to have reinforced Mr. Shawver and other teachers' expectation that science provides different types of learning opportunities around similar concepts and skills.

At Central Elementary School, science instruction did not seem to have established priority among teachers because there was no monitoring of the classroom teachers who were teaching science. Both Mr. Berry and Mrs. Green reported that although they wished to coordinate lessons, it did not happen as much as they hoped. Rather than discussing concepts students were learning, the teachers at Central Elementary School were required to go over skills students missed on DC BAS. Mr. Berry came across messages that Central Elementary School's students were weak on informational texts according to DC BAS results, and teachers were required to cover the reading and math skills on DC BAS. Hence, Mr. Berry said, "the whole thing is, I want them to be able to read science materials and answer questions, that's going to help them with the DC CAS. That's my contribution to helping them with the DC CAS"

(Interview on April 18, 2008). Existing communication among teachers seems to have shaped the types of instructional messages teachers receive.

Third, the existing practice of using assessment in the school seems to have made a difference in how the science teachers experienced the policy pressures associated with math and reading tests. Jefferson Elementary School had a long history of using assessment data to monitor students‘ academic progress and to provide extra support to the students who needed it. The teachers at Jefferson Elementary School looked at the results of DC BAS, but they did not use it to guide their instruction. At Central Elementary School, teachers used DC BAS results as the way to guide their instruction, and test preparation was a school-wide commitment, assigning all teachers to assist. The teachers at Central Elementary School received strong policy pressures that carried the messages that they must teach the standards that are tested. Mr. Berry and Mrs. Green both taught for tested standards.

Summing up, the organizational contexts seem to have shaped the way science teachers incorporated policy pressures associated with the math and reading test by mediating the policy pressures the teachers experienced.

10.2.3 Teacher’s Knowledge

The teachers‘ view of how students learn and existing instructional practices seem to have shaped how teachers incorporated the policy pressures associated with testing. All teachers viewed reading and math skills as being important for learning science. But the teachers varied in terms of how they incorporated policy pressures associated with reading DC CAS, which seemed to have derived from the different instructional practices.

Mr. Berry and Mr. Shawver had similar views about what the goals of school science should be. Both thought that science should inspire curiosity in students about their environment and help them make sense of their environment. Both viewed reading and math skills as being important for learning science. However, they differed about whether instruction should include making students learn scientific investigation. While in Mr. Shawver's class students did scientific investigation by designing experiments and creating hypotheses, in Mr. Berry's class, students were merely introduced to them. Mr. Berry demonstrated the experiments, and students took notes. While Mr. Shawver made sure students understood scientific concepts and vocabularies, Mr. Berry viewed the goal of his science lesson as keeping students interested in science.

Mr. Berry and Mr. Shawver also differed in their view of how to support students who were not strong in reading. Mr. Shawver provided hands-on activities so that students who were not strong readers could participate in scientific discussion, while Mr. Berry viewed reading and writing as prerequisites for students to learn science. Mr. Berry believed his students lacked literacy skills[30]. In addition, Mr. Berry thought students would be able to perform well on the DC CAS if they had good reading skills, even if they did not know scientific concepts. Mr. Berry's view that his students did not have strong literacy skills and his existing instructional practices seem to have made Mr. Berry's instruction more focused on writing. Providing test preparation for reading is in line with his view of what students need to do to learn science.

[30] When I observed his lesson, the students took a long time to copy things written on the blackboard, which Mr. Berry makes students do every lesson. Mr. Berry writes a main question of the activity and draws pictures of experiments on the blackboard, and students copy them. It takes 10 to 15 minutes for students to copy from the blackboard, and he makes sure students copy them. A couple of students write explanations or descriptions by themselves, but for most students copying from the blackboard seemed to be the maximum they can do.

Mrs. Green integrated science into reading and math instruction. Her focus was on the teaching of reading and math skills that students lacked in DC BAS. As Mrs. Green was not familiar with the science standards and she viewed reading as necessary for students to learn science, she taught tested standards of reading by using the science textbook.

10.2.4 Summary

The science teachers rarely came across policy pressures associated with the science DC CAS. The difference among the three teachers derived from the fact that they came across different policy pressures. In addition, teachers‘ views about the role of reading skills shaped different forms of incorporation of reading DC CAS pressures into their science lessons.

In the next chapter, I compare teachers‘ responses between math and science policy pressures and identify the role of high stakes testing for changing teachers‘ classroom instruction.

11.0 ROLE OF HIGH STAKES TESTING FOR CHANGING INSTRUCTION WITHIN THE TEST BASED ACCOUNTAIBLITY POLICY CONTEXT

This chapter presents the analysis of the difference between low and high stakes testing, and the difference between testing and other instructional policy pressures for influencing teachers' instructional change. In the previous chapters, I presented how individual teachers changed their instruction differently in response to math and science DC CAS. I described how the nature of policy pressures, context, and teachers' knowledge shaped individual teachers' instructional change. In this chapter, I discuss the relationship between high stakes testing and instructional change by comparing findings from this study with existing studies. Then, I discuss how policy pressures associated with standards and testing shaped teachers' classroom instruction. My study suggests that although high stakes testing made teachers feel that they had to make their students perform well on the test, and high stakes testing made teachers teach the content being tested, it rarely influenced the pedagogy of instruction; thus, high stakes testing did not lead to substantial instructional change, i.e., change to reform-oriented instruction intended by the district.

11.1 LOW AND HIGH STAKES TESTING AND INSTRUCTIONAL CHANGE

This section compares findings from my study with existing studies by focusing on how the nature of policy pressures mattered in instructional change. My study found that high-stakes testing was more likely to influence teachers' classroom instruction than low-stakes testing. As described in the previous two chapters, math teachers were more likely to change instruction than science teachers in response to DC CAS. All science teachers reported that the testing pressure of science was weaker than the pressure of the math test. As existing studies identified, this perceived weakness of testing pressure derived from the two aspects of policy pressures, pervasiveness (Coburn, 2005) and concreteness of policy messages (Coburn, 2005; Coburn, Pearson & Woulfin, in press; Spillane & Jennings, 1997). My study found that policy pressures teachers received were more concrete in math than science, and the teachers received similar policy messages through multiple policy pressures for math. Thus, the teachers believed the testing pressure was stronger in math than science. However, testing policy pressures were not successful for changing teachers' instruction to reform-oriented instruction.

Policy messages associated with math were more concrete than the ones associated with science. Math DC BAS showed which standards students missed, and this influenced teachers to cover the missed content. Science teachers only had Blue Print, and science teachers did not receive this specific level of information about what they should teach. It is important to note that this level of concreteness of policy message carried by math DC CAS did not provide information about *how* to teach specific content. As described by Kauffman and his colleagues' (2002) study on new teachers' understanding of Massachusetts' standards and assessment

(Kauffman, Johnson, Kardos, Liu & Peske, 2002), the teachers reported that these messages did not tell them *how* to teach this content in their classrooms.

Studies of policy instrument (McDonnell, 2004) suggest that the failure of math and science tests to change teachers‘ instruction derived from the nature of assessment policy. Based on the comparison of several policy instruments, McDonnell (2004) argued that the success of assessment policy depends on whether policy targets (e.g. teachers and parents) could interpret policy messages as policy makers intended. McDonnell (2004) categorized assessment policy as a hortatory instrument because an assessment policy aims to change teachers‘ behavior by making information (i.e., test results) appeal to the values already held by policy targets. According to McDonnell (2004), assessment policy is intended to change teachers‘ classroom instruction by making teachers feel they must improve academic performance by making test results visible. McDonnell (2004) argues that the hortatory policy is based on the assumption that information will appeal to policy targets, and policy targets have the right set of values to take right action.

My study showed that high stakes testing seems to have appealed to teachers‘ sense that they are responsible for improving students‘ academic performance, but the policy did not change their method of teaching to reform-oriented instruction. This finding indicates that DC‘s assessment policy also relied on teachers‘ interpretation of the policy, and that teachers‘ interpretations varied; thus, the policy was not sufficient in changing the core of instruction. In addition, as McDonnell (2004) found in the assessment policy in North Carolina and Kentucky (McDonnell, 2004), my study also revealed that assessment policy itself did not carry sufficient information to guide teachers in changing their classroom instruction. Only the teachers who

had knowledge of reform-oriented instruction (Mrs. Canright and Mrs. Caracelli) could interpret this policy message as it was intended.

In line with existing studies, my study found that policy pressures that were pervasive made teachers change their instruction to cover the content tested on the test. Coburn (2005) found that teachers were more likely to change their classroom instruction if those teachers were connected to a set of policy messages in interlocking and overlapping ways. The policy messages about math instruction and testing were far more pervasive than in science. Math teachers received messages about what they should teach through multiple policy pressures. The Blue Print presented the standards that would appear on DC CAS; the district's professional development taught teachers how to use the Blue Print; and DC BAS results revealed which standards students missed. These policy pressures all informed teachers that they had to cover the content tested on DC CAS. High stakes math testing was more likely to change teachers' instruction because the policy pressures associated with the test presented similar policy messages over time through multiple methods, while in science teachers received policy messages about what to teach sporadically and through far fewer channels.

Although policy messages were more pervasive in math, there was no substantial instructional change, a change of pedagogical approach, among math teachers. This finding was also in line with Coburn's (2005) findings among reading teachers that teachers tend to change instruction in substantial ways when teachers have opportunities to engage with a policy message in sustained and iterative ways. In my study, the teachers had few opportunities to work on DC BAS problems and other test-related materials that would help them re-think their classroom instruction.

11.2 INSTRUCTIONAL CHANGE BY HIGH STAKES TESTING

Instructional change in my study was in line with the findings from existing studies about the influence of teachers‘ existing instructional practices on teachers‘ understanding and response to policy messages (Coburn, 2004; Spillane, Reiser & Reimer, 2002). Although high stakes testing made teachers cover the content, teachers taught the content in the same ways they had been covering content before. Furthermore, teachers‘ existing instruction and belief of how students learn math influenced teachers‘ understanding of policy messages. Below, I present how my study was in line with existing studies about the role of teachers‘ existing practices in shaping their response to instructional policy.

The teachers rarely noticed policy messages that challenged their existing pedagogy, which is in line with existing studies on teachers‘ response to instructional policy (Coburn, 2004; Cohen, 1990; Spillane, Reiser & Reimer, 2002). Only a teacher who was teaching in line with reform-oriented instruction noticed the possibility for instructional change in the policy pressure. While other teachers viewed constructed response problems as a specific format of test problem, only Mrs. Canright noticed the constructed problem touched upon one of the core elements of reform-oriented math, i.e., making students represent math concepts by using various methods, such as sentence diagrams, to understand that math concepts exist in everyday life. Among the six math teachers I interviewed, only Mrs. Canright reported that the instructional change she made in response to DC CAS expanded her view of math instruction. She found that constructed response influenced a positive change in her teaching because she thought it was important for students to communicate mathematical concepts and operations, and constructed response problems push her to teach them. This view of mathematics was in line with her view of good

mathematics instruction. As it is, the policy pressures associated with DC CAS, which was intended to promote reform-oriented instruction, reinforced her view of good math instruction. Other teachers who had been teaching math through teacher-centered instruction did not notice this type of message when they were encouraged to teach for constructed response. Their understanding was that they had to teach the content that appeared in the constructed response.

Although the policy pressures associated with high stakes testing carried the message that there would be consequences if students‘ academic performance did not improve, teachers‘ views of how students should learn math were strong enough to override policy messages that were intended to improve test scores. Mrs. Sommers at Central Elementary School wanted to learn how to teach constructed response for math, and she attended the voluntary professional development session. However, she decided not to teach constructed response to her students. She believed that students should learn skills first. Even though she had heard at the professional development that teachers should encourage students to write answers on constructed response sections because students would be able to get some credit for incomplete answers, she decided to focus on preparing students for multiple-choice questions. At Jefferson Elementary School, Mrs. Canright invited other teachers to look at materials she received from the professional development on constructed response; however, no teacher responded to Mrs. Canright. While Mrs. Canright viewed constructed response as making a positive change in her instruction, other teachers did not view it in this way. The teachers did not think that they needed to incorporate new techniques to teach constructed response. Consequently, most teachers taught math with traditional instruction, while one of the intentions assumed by the DC CAS was to encourage teachers to do reform-oriented instruction.

Further evidence of teachers‘ existing practices and their beliefs about how students should learn as strong determinants of teachers‘ responses to assessment policy is that most teachers in my study took similar approaches to get their students ready for both SAT 9 and DC CAS. Both SAT 9 and DC CAS were high stakes tests, and all teachers reported that that SAT was not in line with the curriculum because it covers content that that classroom instruction did not cover. Although the teachers thought that test preparation for the SAT 9 was separate from their existing classroom instruction, most teachers reported that they made students get ready for the test by covering content that appeared on the test, which is the same approach to test preparation taken for DC CAS. However, teachers varied in their views of how best to prepare students for the test, and their views rarely changed from preparation for SAT 9 to their preparation for DC CAS. The teachers also varied in the way they taught the content for test preparation, which reflected their views about how students learn. For example, every day Mrs. Reed at Jefferson Elementary School continued to have fifteen minutes of drills in which students worked on multiple-choice problems. She had been taking this approach even before SAT 9 began because she thought familiarizing students with the test situation was a year-long process. Mrs. Sommers taught shortcuts and encouraged students to use shortcuts on the test. She was afraid that students would run out of time, and she thought providing strategies to save time was the way to improve test scores. Teaching shortcuts and training students to be able to retrieve a right problem solving strategy seems to be in line with her views of good math instruction. Mrs. Sommers believed each problem has one right strategy to solve the problem and that students should be able to identify the right method. After teaching the concepts behind computation, Mrs. Sommers quickly moved to practice so that students could solve problems quickly. Mrs. Caracelli thought it did not make sense for her to teach shortcuts because she

thought students should learn math by trying different strategies until it makes sense to them. She encouraged students to use the approach that makes sense of them, regardless of time.

Although DC CAS did not change teachers‘ beliefs about good instruction in my study, policy pressures associated with DC CAS did provide some opportunity for teachers to re-think their instruction. For example, at Jefferson Elementary School, the principal asked teachers to discuss the DC BAS results, and the teachers decided to ask students about their thinking processes. At Jefferson Elementary School, the Math Specialist organized a meeting in which teachers discussed how they might want to teach constructed responses. In these meetings teachers had opportunities to re-examine their instruction methodologies. However, as previous studies showed (Coburn, 2004; Cohen, 1990; Spillane, Reiser & Reimer, 2002), teachers often taught by using their existing approach; hence, these opportunities did not result in changing teachers‘ actual instruction practices. Mrs. Reed at Jefferson Elementary School is a case in point. In the grade level meeting, Mrs. Reed and other teachers agreed to ask students how they understood DC BAS problems to find out why students made certain mistakes. She began her lesson by asking students how they had thought about DC BAS problems. However, the pattern of interaction between Mrs. Reed and the students soon changed from Mrs. Reed making students talk about their thoughts to Mrs. Reed explaining a right way of solving problems, the interaction pattern dominant in Mrs. Reed‘s classroom lesson because Mrs. Reed thinks the best way to teach students math is to break down procedures and make students repeat it until students feel comfortable about those specific strategies. This finding is in line with Firestone and colleagues‘ observation of teachers‘ classroom instruction in response to New Jersey‘s low – stakes testing. They found that although the low stakes testing made teachers use math *problems* that promote reform-oriented instruction, the way teachers presented and facilitated

students‘ work on those problems was not in line with the reform-oriented instruction. As McDonnell‘s (2004) conceptualization of assessment policy suggests, teachers‘ knowledge and capacity shaped the success of high stakes testing policy.

As discussed by McDonnell (2004), the assessment policy seemed to have appealed to the values teachers held. All of the six math teachers I interviewed reported that they were concerned about students‘ performance for both SAT 9 and DC CAS, and that they did test preparation. The teachers also reported that it was their responsibility to improve students‘ performance on these tests. For example, Mrs. Caracelli reported that she believed that the improvement on SAT 9 and DC CAS reflected her capability as a teacher. Mrs. Green reported that she had felt sorry for her students when they fell ―below basic‖ on the test, and she felt embarrassed when the score of her students on SAT 9 was lower than other classes‘ scores. These examples indicate that the test results made teachers feel that they had to improve students‘ performance. To do so, all teachers taught the content that was covered by the test.

In summary, high stakes testing policy did not influence teachers‘ existing practices and beliefs about good instruction—this finding is in line with Diamond‘s (2007) finding that high stakes testing rarely influences pedagogy. Although the high stakes testing was supposed to make teachers change instructional methods by providing incentives and by appealing to teachers‘ values (McDonnel, 2004; Spouvitz, 2009), the policy was not sufficient for changing their pedagogy. My study found that even the assessment that was in line with the standards rarely influenced the pedagogy of instruction. Although teachers noticed that SAT 9 and DC CAS were different tests, they rarely noticed the differences in instructional approach that DC CAS was supposed to convey. Consequently, the teachers varied in their instruction, and most of them continued with traditional instruction.

11.3 STANDARDS, CURRICULUM, AND ASSESSMENT'S INFLUENCE IN CHANGING TEACHERS' CLASSROOM INSTRUCTION

In spite of the fact that congruence is a central part of the theory of standards-based reform, the standards, curriculum and testing did not carry a uniform instructional policy message when teachers came across them. The standards, curriculum, and assessment differently influenced teachers‘ classroom instruction. In this section, I discuss how teachers differently responded to policy pressures associated with standards, curriculum, and assessment. By doing so, I discuss the role of high stakes testing in relation to standards and curriculum in changing teachers‘ instruction to be more reform-oriented. As Diamond‘s (2007) study found, I argue that high stakes testing narrowed instruction for most of the teachers because it promoted different conceptions of good instruction from standards and curriculum. In addition, I claim that while the prescribed curriculum could change the activities teachers introduced in the classroom, the curriculum itself did not change teachers‘ views of how students learn math or change their beliefs about good math instruction. Below, I describe how policy pressures associated with standards and curriculum differently influenced teachers‘ instruction.

The curriculum, *Everyday Mathematics*, influenced teachers‘ instruction to varying degrees, which sometimes went beyond influencing the content. For two teachers who were inclined to teach in reform-oriented ways, *Everyday Mathematics* extended this type of instruction and reinforced the instructional practices and beliefs that reform-oriented instruction makes students learn math. For the rest of four teachers, *Everyday Mathematics* seemed to get them to introduce math problems that appeared on Math Journal, a student workbook that was to be used in the lesson. At Jefferson Elementary School, all teachers used *Everyday Mathematics*

as the central part of their classroom instruction. Everyday Math was the curriculum three out of four teachers closely followed in their lessons, and all teachers at Jefferson Elementary School reported that *Everyday Mathematics* makes students learn math better than previous curricula. At Central Elementary School, Mrs. Green and Mrs. Sommers also tried to use *Everyday Mathematics*; however, they often supplemented it with other materials. Although the teachers varied in their understanding of what types of instructional change Everyday Math required, most teachers tried to follow the curriculum. The teachers at Central Elementary School incorporated part of the activities in the Math Journal, and most teachers at Jefferson Elementary School taught the Math Journal. Teachers who were inclined to reform-oriented instruction, Mrs. Canright and Mrs. Caracelli at Jefferson Elementary School and Mrs. Green at Central Elementary School, noticed that the approach of instruction was different, while other teachers who focused on the mastery of problem-solving procedures did not notice that *Everyday Mathematics* required teachers to change their conceptions of how students learn math.

Everyday Mathematics influenced classroom instruction at the level of making students work on a complex math problems, while DC CAS made most teachers cover tested skills. All teachers noticed that *Everyday Mathematics* is different from the previous curriculum because they were told that teachers should move from one concept to another sooner than in the previous curriculum. Because the district mandated professional development when it was implemented, all teachers knew about the spiral structure of *Everyday Mathematics*. All teachers noticed that the order and pacing of lessons were different from their previous curriculum. All teachers noticed that the curriculum brought in new math activities, new order of lessons, and faster pacing, rather than focusing on a specific skill or math concept. For teachers at Jefferson Elementary School, Everyday Math also made students work on problems that required the use

of multiple strategies because the teachers covered the problems on the Math Journal. Teachers made students work on math tasks that were not limited to simple computation practices. As it is, *Everyday Mathematics* at least brought in extended math tasks to the students.

However, how teachers taught math problems prescribed in the Math Journal varied; thus, the instructional change *Everyday Mathematics* produced in teachers varied. While some teachers, for example Mrs. Canright and to a lesser degree Mrs. Caracelli, used problems and activities suggested in *Everyday Mathematics* by allowing students to discuss their problem solving processes, other teachers used problems on the Math Journal, but often omitted discussion among students. These teachers reviewed math problems from the Math Journal by making sure that students mastered specific problem-solving procedures. These teachers were teaching new problems but by using pre-existing teacher-centered instructional approaches. Although all teachers at Jefferson Elementary School told students that one problem could be solved through different approaches, the majority of teachers reported that their interaction pattern with students was the same as with the previous curriculum; thus, these teachers taught new types of math problems by using old strategies. While most teachers at Jefferson Elementary School reported that *Everyday Mathematics* changed their instruction, the change of instruction was introducing new problems and letting students work on problems that called for multiple strategies. The teachers did not seem to have changed the degree to which they allowed discussion among students and the degree to which they guided students‘ exploration when they asked questions. Furthermore, teachers‘ view of how students learn math did not seem to change. For example Mrs. Reed and Mrs. Miller at Jefferson Elementary School said that they believed making students get used to problem-solving procedures was the way for students to learn math, which is the approach Mrs. Reed and Mrs. Miller took when they used the

previous curriculum. Although the district intended for Everyday Math to change the instruction approaches of teachers, the change was limited. Only the teacher who believed in the reform-oriented instruction in the first place noticed the difference in approaches to math instruction between Everyday Math and the previous curriculum. Furthermore, the most reform-oriented teacher in my study, Mrs. Canright, did not necessarily follow the order and problems provided by *Everyday Mathematics*. She incorporated tasks that *Everyday Mathematics* envisions, but they were not necessarily from the Math Journal. One teacher at Central Elementary School was least influenced by Everyday Math.

The standards were even less likely to change teachers‘ instruction than *Everyday Mathematics*. The policy pressures associated with the standards carried the message that teachers should plan math lessons around power standards. Most teachers at Jefferson Elementary School followed *Everyday Mathematics* rather than creating lessons around power standards. Although the teachers at Central Elementary School identified power standards and they planned instruction to cover the power standards, the teachers did not create standards-based worksheets that included power standards, engaging scenarios, or performance assessment.

The comparison between policy pressures associated with *Everyday Mathematics* and standards suggests that the teachers tended to incorporate policy pressures that were concrete, such as *Everyday Mathematics* that came with the Math Journal. Although all teachers agreed with the idea that not all standards were equally important, that identifying power standards and supporting standards helps teachers to plan instruction, and that students‘ performance should be measured in multiple ways, no teachers actually used standards-based worksheets to plan as they were taught in the professional development.

In my study, standards tended to influence teachers‘ instruction indirectly. Policy pressures associated with the standards were repeated district-wide professional development that taught teachers how to plan instruction based on the standards. The impact on actual classroom practices was weak compared to *Everyday Mathematics*. Only one teacher reported that science standards changed the content of her science lesson. Until then, she created her own report card based on the FOSS kit. When she attended a series of professional development sessions on the standards from April 2005 to summer of 2006, she consulted with the principal and she began to teach the content that the report card states by supplementing with other resources. She reported that she had begun to feel she had to teach for the standards by seeing and experiencing many training sessions on standards. As a result, she taught the content that appeared on the science standards by using the FOSS kit as much as possible. No teachers at Jefferson Elementary School reported they created lesson plans around power standards. For three other teachers, the professional development on standards provided opportunities to discuss how the concepts and skills taught in one subject may relate to other subjects; professional development on standards also made teachers see that teaching for the standards is important. For example, at Jefferson Elementary School, Mr. Shawver reported that unwrapping the standards and looking for vertical alignment might have made him pay more attention to the concepts and vocabulary students were learning in other academic subjects. As I described, the teachers followed *Everyday Mathematics*. The teachers at Central Elementary School used the term power standards until 2008. Both Mrs. Reed and Mrs. Green reported that they pay attention to power standards; however, they did not create the standards-based workbook. As it is, they used the power standards to check the content and how much time they should spent in relation to non-power standards.

Only one aspect of the standards directly influenced teachers‘ instruction. All teachers reported that they paid attention to the topics listed on the report card Although most teachers were skeptical about whether the topics listed on the report cards reflected power standards, teachers covered the topics because, they said, the parents would look at the report card. The teachers varied in the degree to which they planned instruction around the topics on the report card. Mrs. Sommers at Central Elementary School reported that the topics on the report card should have reflected the power standards, and she created classroom assessment centered around the topics that appeared on the report card. Other teachers reported that they made sure to cover the topics on the report card. The influence of the report card was less than the influence of DC CAS because it was up to teachers how much time they spent on topics that appeared on the report card. Furthermore, each teacher decided what level of performance met with the standards. Thus, rather than making teachers rethink their classroom instruction, the influence of the report card provided a checklist of content to be covered.

High stakes testing changed the content of instruction because teachers incorporated frequently tested standards into their instruction. The degree of influence of DC CAS was stronger than the report card. When teachers found the gap between the content covered by standards, *Everyday Mathematics* and DC CAS, they made sure that the content that would appear on DC CAS was covered. As Watanabe (2007) found in the standards and high stakes testing in North Carolina (Watanabe, 2007) and as Diamond (2007) found in the interviews with teachers in Chicago, policy pressures that teachers received associated with DC CAS boiled down to one simplified interpretation: that teachers should cover the tested content.

Summing up, neither policy pressures associated with curriculum and standards, nor tests changed teachers‘ views of good instruction as intended. Rather, the teachers taught the content

suggested and introduced activities by using their existing instructional practices. Although standards, curriculum, and DC CAS were meant to promote reform-oriented math instruction, the teachers‘ instruction seldom changed to this direction.

11.4 CHANGE OF INSTRUCTIONAL PRACTICE IN STANDARDS-BASED AND TEST-BASED ACCOUNTABILITY CONTEXT

My study asked teachers about their instructional change since they began teaching. The data suggests that policy pressures associated with test and standards were not sufficient for changing teachers‘ views of good instruction. Among a total of eight teachers, two teachers, Mrs. Canright and Mrs. Green, reported a change in their views of good instruction and a change in patterns of interaction with students in the classroom. These types of instructional changes were accommodation, according to the categorization by Coburn (2004), because it changed belief about and patterns of interaction. In this section, I present how this substantial instructional change happened to argue that policy pressure from standards, curriculum, and assessment was not enough to lead to substantial instructional change. As Coburn (2001) found, having support from colleagues who could closely work with teachers was necessary for teachers to transform their pedagogy of instruction.

Mrs. Canright and Mrs. Green‘s experiences indicate that instructional change is shaped by many factors, rather than by the policy pressures from high stakes testing. Their experience also indicates that the change of pedagogy is a long process. The change of curriculum, standards, and assessment were factors that triggered the instructional change. Their accounts of

instructional change indicated there had been many factors that shaped the change of their instruction. As studies on instructional change suggest, these factors are teachers‘ interests, their own experience as learners of math (Spillane, 2000), their opportunities to reflect their own practices, their understanding of student learning, and their opportunities to see other teachers‘ practices (Coburn, 2001, 2004). By utilizing these opportunities, Mrs. Canright and Mrs. Green tried to change their practices and they re-conceptualized their own approach. Policy pressures associated with high stakes testing was one of the opportunities one teacher, Mrs. Canright, took to expand her repertoire of instructional strategies, and the high stakes testing reinforced her reform-oriented instructional approach.

In contrast, for Mrs. Green, policy pressure associated with high stakes testing inhibited her attempt to change her instruction to reform-oriented instruction. Mrs. Green reported her teaching had changed because the district introduced Everyday Math. She also reported that during that time she had reflected on how she learned mathematics, made connections with her educational philosophy, began reading the teacher‘s manual, and started trying some lessons from Everyday Math. She saw students were understanding math better. She reported that all these incidents made her shift her instruction from traditional approach to ―inquiry-based" instruction. Although her teaching from Everyday Math did not happen as often as she wanted it to, and she was aware that she often used direct instruction, she saw her instruction change when she began teaching with Everyday Math.

Mrs. Canright‘s understanding of how special education students learn is a substantial change because the change of conceptualization made her change how she supports special education students and what types of tasks she provides them. She compares her old views and current views in this way. She reported that before she had thought differentiation was providing

easy work for students with special needs. She also reported that she now believes she could hold special education students accountable at a higher level of work. She reported that when she had begun teaching she did not know much about how special education students learn. She worked closely with a special education teacher who worked in her classroom. Mrs. Canright began to see her students working on the same level with regular students in math. At the same time, she learned various strategies to teach different age groups by helping at a pre-school part time. She reported that all these experiences helped her to uncover more about how students learn.

Both teachers said that the change of instruction happened over time as they worked with other teachers, by seeing how new instruction plays out in their classroom through close consultation with their colleagues as described in previous studies on teachers‘ instructional change (Coburn, 2001). High stakes testing played a role in this process. Both teachers had been under pressure to make students perform better, and they were looking for ways. However, the pressure from high stakes testing to make students perform well was merely a lever. Information about how to change instruction came from multiple sources, such as own learning experiences and working with colleagues for sustained time.

11.5 SUMMARY

This chapter analyzed the impact of policy pressures associated with high stakes testing, curriculum, and standards on teachers‘ classroom instruction. Policy pressures associated with standards, curriculum, and high stakes testing did not convey uniform instructional messages to

the teachers. The teachers selected the policy pressures they incorporated. High stakes testing influenced the content of instruction. Curriculum and standards differently influenced teachers' classroom instruction. Overall, the teachers tended to notice instructional messages about content, skills, and the order they should teach in, but only the teachers who were inclined toward reform-oriented instruction before noticed messages about instructional change. The change of approaches in instruction happened when teachers came across multiple policy pressures that originated not only from the district, such as high stakes testing, standards, and curriculum, but also when they saw their colleagues teaching in new ways.

12.0 SUMMARY AND CONCLUSION

My study examined the role of high stakes testing policy by drawing upon concepts from New Institutional Theory and Cognitive Approach. I conceptualized my study as in line with studies that examine individual responses to institutional pressures that promoted new ideas and practices (Coburn, 2001, 2005). I also designed this study to expand our understanding about the influence of high stakes testing on teachers‘ classroom instruction. Below I discuss how my data was in line with current discussion in New Institutional Theory and in line with prior understandings of the role of high stakes testing in classroom instruction. Then, I discuss areas for further research.

12.1 THIS STUDY AND NEW INSTITUTIONAL THEORY

In line with current new institutional theory discussion (Coburn, 2004; Oliver, 1997; Scott, 2003), my study showed that the interpretation of institutional logics influenced how institutional logics spread among individuals. My study found that policy pressures associated with high stakes testing were the carriers of institutional logics about good instruction and about how assessment should be used. In line with other studies (Coburn, 2001; 2005), my study also found that some institutional pressures reached teachers and that teachers selectively incorporated

policy messages. The study showed that teachers‘ existing practices and beliefs influenced whether or not teachers noticed the messages and how they responded to these messages.

This finding made the following contribution to the institutional theory discussion. My study documented the change in the messages as the policy pressures moved from the district and school to teachers by focusing on the nature of policy pressures, i.e., carriers of institutional pressures. The role of the institutional carrier is one of the areas the institutional theorists have studied (Coburn, 2001, 2005; Spillane & Jennings, 1997, Scott, 2003); Scott (2003) categorized types of careers of institutional pressures into the following groups: a) symbolic systems, b) relational systems, c) routine, and d) artifacts, and he argued the need for examining how different types of carriers matter in the spread of institutional logics. My study showed that a relational system was important when artifact was used. The teachers received different policy messages from the DC BAS results depending on how the principals facilitated the training at each school. While teachers at Jefferson Elementary School learned how to read the DC BAS reports, teachers at Central Elementary School understood that they had to teach for the standards that were tested on DC BAS. This difference was important because while teachers at Central Elementary School thought they had to teach certain standards, some teachers at Jefferson Elementary School began to ask students about their thinking, which could trigger changes in instructional approach. A principal, a person who has authority over a teacher, articulated institutional logics by making teachers do certain activities that resulted in articulating the meaning of institutional logics; thus, the relational system mediated policy messages.

The studies of education policy have documented the roles that principal and school leaders, as the mediators of policy messages, take on within accountability policy contexts (Anagnostpoulos & Rutledge, 2007; Diamond & Spillane, 2004). These studies stated that

principals could utilize resources and tools to articulate the school‘s problems and describe steps the school should take to solve those problems. These studies documented the importance of a principal‘s role in articulating what high stakes accountability policy means for teachers‘ work. My study also showed this mediating role of the principal.

This finding further solidifies the assumption that schools are not decoupled from accountability policy pressures (Elmore, Abelmann & Fuhrman, 1996; Elmore, 1999-2000) by showing accountability policy pressure made teachers pay attention to testing, but actual instructional change varied. Educational researchers have argued that school must have been responding to improve academic performance that is promoted by accountability measures, and schools are not buffered from policy pressures. My study found that accountability pressure in DC included two logics about instruction: one is teaching for the test and another is changing classroom to reform-oriented instruction. My study found the first logic was reaching teachers but the second logic varied. Thus, my study showed accountability policy is not uniform, and although the pressure to improve student academic performance spread among teachers and although the teachers changed the content of instruction, teachers rarely changed pedagogy.

Further evidence that delineate institutional theory discussion is the role of perceived legitimacy in individual‘s response. Teachers‘ perceptions about to whom and to what they are accountable were important factors in understanding teachers‘ responses to policy pressures. However, teachers‘ responses did not cover the depth of the institutional logic. All teachers incorporated the content listed in the report card, and all teachers said that this was because parents would see the report card. All teachers felt they had to improve student performance on DC CAS. However, teachers varied in the degree of change of instruction when they responded to policy pressures of the report card and DC CAS. Some teachers covered the content listed on

the report card just to be able to say that they had covered the content. Other teachers tried to fit in the content in the right proportion of time so that the content listed on the report card appeared to be important content. Similarly, while all teachers felt they must improve students' academic performance, their instructional change varied. The evidence from my study suggests that Oliver's theory of organizational response, "Lower the degree of social legitimacy perceived to be attainable from conformity to institutional pressures, the greater the likelihood of organizational resistance to institutional pressures" (Oliver, 2003, p. 160), is applicable for this individual case, because teachers' perceptions of parents as legitimate feedback sources made them change their instruction, and the teachers' views of DC CAS performance as something they had to comply with made them teach for the test. At the same time, the change in individual behavior varied, which suggests that the perception of social legitimacy did not fully shape how they incorporated institutional pressures. Teachers' knowledge and beliefs about learning shaped their responses to DC CAS.

Summing up, my study was in line with the current institutional theory discussion to understand the process through which individuals respond to institutional pressures. The study presented the role of carriers and existence of conflicting institutional logics within the accountability policy pressure. The study presented that while the pressure to improve academic performance reached teachers, institutional pressures of good instruction varied in changing teachers' behavior.

12.2 THIS STUDY AND STUDIES OF THE IMPACT OF HIGH STAKES TESTING ON CLASSROOM INSTRUCTION

My study presented evidence that delineates our understanding of the role of high stakes testing on teachers' instruction in the following ways: First, my study presented the influence of high stakes testing on teachers' classroom instruction in relation to other policy instruments, such as standards and curriculum. Although the standards-based reform aims at improving students' academic performance by aligning the standards, curriculum and assessment review of policy pressures associated with these components revealed that they promoted different concepts of how classroom instruction should be implemented. In line with Watanabe's (2007) study, policy pressures associated with tests tended to narrow the messages about instruction to the content of instruction because the teachers receive materials and information about *what* teachers should teach. This narrowing of instruction was promoted not only by giving Teacher Report for DC BAS and Blue Print to the teachers, but also by the way these materials were presented to the teachers. By using individual teachers' encounters with policy pressures associated with the test as a unit of analysis, the study presented what types of policy pressures teachers noticed and how they incorporated those pressures.

Second, concepts from New Institutional Theory helped to document what high stakes testing policy entails and how high stakes testing pressure is different from low stakes testing pressure. The study showed that policy pressures associated with high stakes testing carried a message that teachers must improve student academic performance, and this type of policy pressure was not sufficient for guiding teachers' instructional change. Although some policy pressures provided information about *what* teachers should teach, most of the time, it was up to

teachers *how* they should teach it. The difference between high stakes testing and low stakes testing in DC's public schools was the frequency and nature of policy pressures. My finding that teachers reported low testing did not change their instruction was different from findings from studies that reported that low stakes testing makes teachers teach for the test. Possible reasons for this difference is that the policy messages about the science test were extremely few, and it was a new test, while other studies that reported the influence of low stakes testing pressure studied testing policy that had been in place for some time; thus, teachers had received more policy pressures (Firestone et al, Firestone, Montils & Schorr, 2004; Smith, 1991).

Third, my study adds data on how school differences influence teachers' responses to highs takes testing by collecting data from two schools. Existing studies found that a school's organizational capacity and norms make a difference in school-level responses to high stakes testing pressure (Carnoy, Elmore & Siskin, 2003, Lipman, 2004). My study documented the principal's ability to articulate testing policy, teachers' views of good instruction, available resources to support instruction, and teachers' and school leaders' understanding of the meaning of assessment data and experience using them. The study also showed the importance of existing instructional practice and instructional support as a way of continuing with successful instruction. In line with the studies that described how a teacher's relationships with other people in the school shape their instructional decision (Coburn, 2001, Diamond, 2007), my study also showed that a school that lacked this type of organizational capacity and knowledge tended to narrow their instruction.

12.3 AREAS OF FURTHER STUDIES

My study examined how eight teachers differently changed classroom instruction in response to the district assessment that was intended to be in line with the standards and curriculum to promote reform-oriented math instruction. I found all teachers were concerned about the test; however, the test did not change pedagogy. Out of the six math teachers, only one teacher was teaching reform-oriented math; however, the high stakes testing had not been the condition for changing this teacher's instruction. Although high stakes testing policy is based on the assumption that incentives will make teachers teach for standards and improve teaching and learning at schools, my study found this was not the case. Rather, as a recent review of high stakes testing by Supovitz concludes, my study also showed it was a "weak intervention" for instructional change (Supovitz, 2009, p.211) because it did not change pedagogy and not all teachers incorporated the suggestions of instructional change. Most teachers continued with instructional approaches that were not in line with reform-oriented instruction.

Based on my study, I suggest that the following further studies are needed to better develop concepts in institutional theory and to better conceptualize the role of high stakes testing in changing classroom instruction. First, there is a need for data on other types of institutional carriers and how a combination of particular carriers and the nature of institutional logics shape individuals' responses. My study identified the relational system as being an important carrier in the context of DC public schools. As interviews were my main data collection method and I did not do sustained observation, my study did not examine how routines and artifacts carry and modify the policy messages examined. In addition, there is a need for more cases that examine the role of carriers in different types of organizational environments and different institutional

logics. My study analyzed the role of carriers in the organizational environment where conflicting institutional logics existed and where policy pressures did not promote a single institutional logic. For example, after I had collected data, DCPS introduced a framework of instruction called IMPACT that is tied to teacher evaluations. This policy requires teachers to teach and interact with students in a certain way, and teachers are to be periodically monitored to determine whether they are making changes in their instruction. This policy combines strong monitoring and less space for interpretation; thus, IMPACT provides a comparative case to understand the role of carriers in the change of individual behavior and belief.

Second, in order to improve our conceptualization of the role of high stakes testing on teachers‘ classroom instruction, more studies on different types of relationships between the school and the district, between different types of schools, and between districts with different types of standards, curriculum, and high stakes testing are needed. The two schools I collected data from were very different in terms of their relationship with institutional environment and organizational norms and capacities. Jefferson Elementary School was almost buffered from the district‘s institutional policy, while Central Elementary School was not. Jefferson Elementary School had a shared norm about the school‘s instructional initiative, while Central Elementary School did not. My study presented these differences as shaping how individual teachers changed instruction; however, the study did not indicate what exactly these norms are that positively shape teachers‘ instructional change, and how much the school should be buffered from the district‘s institutional pressure for sustaining the school‘s shared norms about instructional initiative. In order to understand how the internal workings of schools shape teachers‘ instructional change, there is a need for more studies with schools and teachers that are theoretically sample-based. For example, in order to articulate how different degrees of

institutional pressure influence teachers‘ classroom instruction, a researcher may want to select teachers who have been teaching reform-oriented instruction for a while from three schools. Two schools are similar in the degree of institutional pressures from the district. Having three schools helps to identify if the degree of institutional pressure associated with assessment influence teachers‘ classroom instruction. Another example is studying teachers who are implementing district curriculum with fidelity. This sampling will provide a detailed description of what component of curriculum teachers tend to change in response to high stakes testing. Cross-case analyses of these studies will help us better conceptualize factors that mediate institutional pressure.

Summing up, the study examined how the process of high stakes testing shape teachers‘ instruction. To refine our understanding, it would be beneficial to use other forms of data collection and conduct more case studies.

APPENDIX A

INTERVIEW PROTOCOL FOR DISTRICT OFFICIALS

Purpose

The purpose of this interview is to acquire information about the following issues:

1) How did instructional approaches change in the DC Public Schools when the new standards and assessment, DC CAS, were implemented?
2) How were new instructional messages carried to the teachers?

Based on the time table of instructional policies and implementation process I created from the document review, I modified the questions based on their experiences and involvement in creating and implementing the policies.

CHANGE OF POLICY How was the new standard different from previous standards?

1) How was the new standards developed?
- Who were involved?
- What were the issues discussed?

2) How were the new standards different from the old standards?
- Good math/science instruction
- How students should learn
- Teacher's role
- Curriculum
- Assessment
- Anything else?

3) How did teachers and schools come to learn about the new standards?
- What types of professional development was provided? Who provided? What teachers do? How often?
- Were there other materials used?
- What strategies did the district take to inform teachers and schools? Why?
- Do you have any materials used around that time?

4) What is your observation about the spread of the standards?
- Compared to the previous standards, how the standards spread to schools?

CHANGE OF ASSESSMENT

1) How was the information about SAT 9 delivered to teachers and schools? How were teachers and schools informed about goals, characteristics, and use of the assessment?
   - Were there professional development?
   - Were there other materials?

2) What is your observation about the alignment between SAT 9 and curriculum/standards?
   - Are the descriptions on the newspapers accurate?

3) How was the information about DC CAS delivered to teachers and schools? How were teachers and schools informed about goals, characteristics, and use of the assessment?
   - What types of professional development provided?
   - How did teachers learn how to prepare their students for test?
   - What is the relationship between DC CAS and the standards and curriculum?
   - What materials were available for teachers about teaching for the state test?
   - Did the materials or professional development opportunities change over time?

INSTRUCTIONAL POLICY: SCIENCE

1) What is the curriculum in the district?
   - FOSS kit, New textbook

2) What is the guidance from the district about how to teach science? How is it delivered to teachers?
   - What did professional development look like?

3) What does good science instruction should look like?
   - How students should learn science?
   - What is the role of teachers?

4) What instructional resources are available for teachers to plan and learn about instruction?

OTHER EVENTS/ACTORS

1) Besides the events we talked, were there any other events, organizations, and reports that were important to understand the district's instructional policy?
2) Who else should I talk to learn about the change of district instructional policy and testing policy?

APPENDIX B

INTERVIEW PROTOCOL FOR TEACHERS

Purpose

The purpose of this interview is to acquire information about the following issues:

1) How have teachers encountered and responded to policy pressures?
2) How have teachers' responses been shaped by the nature of policy pressure, context and their existing knowledge and practices?

The questions below are general guiding questions. Follow up questions were asked to acquire concrete descriptions.

Interview Questions

TEACHING HISTORY

1. Walk me through your teaching career:
    - What subject do you teach? How long have you taught in this school? How many years have you taught this subject? How many years have you taught testing grade?
    - Did you teach SAT 9 grades?
    - Did you teach DC CAS grades?

2. How many classes do you teach? What does your daily schedule look like?
    - Is there a school-wide schedule for specific academic subject?

3. What training did you have? When and where did you receive your training?
    - What instructional approaches did you learn?
    - What was your reaction to different instructional approaches?
    - How do they help you make instructional decisions?

GOOD INSTRUCTION AND PRACTICE

1. How do you plan your instruction?
    - What do you use to plan instruction? Why?
    - How did you learn to plan instruction in this way?
    - Did your approach to planning change? If so, how?

2. How did your (math/science) teaching change since you began teaching?

- Why did you change?
- When did it happen?
- How did the change happen?
- How did you learn about it? Who suggested you to do so?

3. What is good math/science instruction?
   - What does it look like in your lesson?
   - How students learn (math/science)?
   - How do students differently learn, if any?
   - What support/assistance do you provide to help students learn?
   - What is your role in the good instruction?
   - How often does the good instruction happen? What makes you to do good instruction? What does not? What are the challenges do you have to do good instruction?

4. Did your view of good instruction and/or how students learn change over time? If so, how did it happen?
   - When did it happen?
   - How did it happen?

5. What your typical lesson looks like?
   - How do you begin the lesson?
   - Do you use Everyday Mathematics/Textbook?
   - What do students do?
   - Do students work with others?
   - What feedback do you provide?

6. In what occasions do you have to change your lesson plan?
   - What shapes your instruction?
   - How often does the mid course change happen?

USE OF ASSESSMENT

1. How do you use assessment?
   - What assessment do you do? (including quiz, homework assignment, benchmark test)
   - How do you use the assessment?
   - How often do you use the assessment?

2. What are the differences among DC CAS, DC BAS, and other assessment?
   - Do you get different types of information?
   - Do you use differently?
   - What do they inform you? How do you use the information?

3. Did your way to use the assessment change over time? If so, how did it happen?

INFLUENCE OF STANDARDS

1. How did the standards shape your instruction?
    - How did you come to know about the old standards? Did you have PD?
    - How did it influence your instruction? What information do you use for planning instruction?
    - What did the professional development on the standards look like? What did you do? What did you learn?
    - How are you using the learning from the professional development in your instruction?

2. How were the new standards different from old standards?
    - What does the old standards look like?
    - How did you learn about the old standards?
    - How did you use the old standards? How did you teach?

INFLUENCE OF STATE TEST

1. How did SAT 9 influence your instruction?
    - Did you change the content?
    - Did you change the assessment you do?
    - Did it change your instructional approach?
    - How did you come to know about old state test, SAT 9? Did you receive training?
    - How did the training influence your instruction?
    - How did you use the information from SAT 9?
    - What information do you use for planning your instruction? How did you learn how to use it?
    - How did the SAT 9 relate to classroom instruction, standards, and curriculum?
    - Do the descriptions on the newspaper reflect your observation and experience?

2. How did DC CAS influence your instruction?
    - Did you change the content?
    - Did you change the assessment you do?
    - Did it change your instructional approach
    - How did you come to know about DC CAS?
    - How do you use them? How does the information from DC CAS influence your instruction? What changes did you make in your instruction?

3. How do you make students ready for DC CAS?
    - How did your approach change?
    - How is your test preparation different from the time you prepare students for SAT 9?

4. How does DC CAS relate to standards and curriculum? How did you learn about it?

5. What other test, assessment, evidence, data do you and your school use? How do you use them? How did the use of evidence/data change?

6. Some people say high stakes test makes teachers teach for test. How do you comment on it?

7. NCLB requires all schools to meet with 100 % proficiency on state test by 2014. Do you think it is obtainable?

INFLUENCE OF CURRICULUM

1. How did the old curriculum shape your instruction?
   - How did you come to know about old curriculum?
   - What did it look like? What did students do? What did you do?
   - Did you use the curriculum?
   - How did it influence your instruction?

2. How does the new curriculum shape your instruction?
   - How did you come to know about new curriculum?
   - How do you use it?
   - Do you adjust curriculum? Why? How did you learn about it?

3. How does the new curriculum differ from the old curriculum?

OTHER POLICY PRESSURES

1. What influencing your instruction?
   - How principal, other teachers, parents influence your instruction?
   - How do students influence your instruction?
   - Is there anything you do not like but doing in your instruction?

2. What materials/resources are available for you to plan instruction? How do you use them?

3. What resources do you wish to have to help you plan and/or teach?

CLOSING

1) Is there anything I left out to understand change of your instruction?
2) Is there anything I left out to understand change of DC's instruction policy, guidance, and assessment?
3) Do you have any materials you mentioned?

After I listed up policy pressures teachers encountered, I created a set of cards that include sources of policy pressures and ~~–p~~edagogy" and ~~–c~~ontent." I asked teachers to sort out

the cards into either ~~p~~pedagogy," ~~c~~ontent," ~~b~~oth," or ~~n~~one." After the teacher finished sorting, I asked her/him to elaborate why and in what way each card is sorted. I did this activity to make sure I covered policy pressures the teacher mentioned and to ask about the influence of policy pressures that were not mentioned by this specific teacher but were mentioned by others.

For Math Instruction

- Students
- Parents
- Principal
- Grade level teachers
- Specialists (Specialist Teachers, Data Specialists, Counselors at school)
- Other teachers (both in school and outside of the school)
- Mentor teachers
- Education Training, Program
- Philosophy
- School initiative
- Own experience
- Report Card
- Old Curriculum
- Current Curriculum
- DC CAS (state assessment)
- DC BAS (benchmark assessment)
- Other Assessments
- Current Standards
- Old Standards
- Professional Development

For Science Instruction

- Students
- Parents
- Principal
- Grade level teachers
- Specialists (Specialist Teachers, Data Specialists, Counselors at school)
- Other teachers (both in school and outside of the school)
- Mentor teachers
- Education Training, Program
- Philosophy
- School initiative
- Own experience
- Report Card
- FOSS kit
- New Textbook
- DC CAS (state assessment)

- Other Assessments, Student Work
- Current Standards
- Old Standards
- Professional Development

APPENDIX C

CLASSROOM OBSERVATION PROTOCOL

Purpose

The purpose of observation of classroom instruction was to acquire concrete example of practice.

I took note of the observation with descriptive system of recording with functional unit and IRE (initiative, response, evaluation) as observation unit (Evertson & Green, 1986).

The observation notes included the following information.

Observation Protocol

Date:
Time:
Place:

- Description of the classroom‘s physical condition (layout, decoration, signs)
- Number of students and seating arrangement
- Topic of the lesson
- Questions teacher ask
- Students‘ response
- Teacher‘s response (evaluation, further questions, support)
- Students‘ task
- Time allocated to the above four points
- Materials present or used

After I type up the notes, I described the characteristics of instruction around the following aspects of instruction.

For Math

1) Nature of tasks teacher provided:
    - Is deep and complex (includes multiple skills)
    - Allows students to connect concepts within math and across subjects,
    - Is relevance to students‘ experiences and knowledge

2) Process of thinking and investigation students experience is
    - Problem solving,
    - Active thinking,
    - Multiple solution strategies,
3) Classroom interaction, discourse is:
    - A teacher encourages students to explain
    - Students critique and explain strategies each other
    - A teacher assists students progress their thinking, helps students clarify concepts
    - A teacher encourages multiple forms of presentation of mathematical ideas

For Science

1) Nature of tasks/activities teacher provided:
    - Is hands on experiences and minds on experiences
    - Helps students integrate math, and make practical applications to technology
    - Is relevant to students‘ experiences and challenge their misconceptions about environment
2) Process of thinking and activities students engaged in is
    - Active process of investigation, including students reflecting own ideas, collecting evidence, making inference and predictions, critical review of evidence.
    - Problem Solving,
    - Replicating important experiments that confirmed knowledge about the natural world in the classroom
    - Extended time of investigation
3) Classroom interaction, discourse is:
    - A teacher encourages students to explain.
    - A teacher encourages students to collaborate, communicate, critique
    - A teacher asks questions challenge students and promote inquiry.
    - A teacher models scientific inquiry.
    - A teacher use multiple data to assess student learning

APPENDIX D

CODING STRUCTURE

Base Code

- **S**chool (AYP status)
- Academic Subject (math, science)
- Training Background (elementary focus, secondary focus, policy context when they received the training, instructional approach taught)
- Years of Experiences

Analytic Code (Codes from theoretical framework and existing studies)

- Teachers‘ responses to policy pressures
    - Rejection,
    - Symbolic response,
    - Parallel structure,
    - Assimilation
    - Accommodation, and
    -
- Teachers‘ pre-existing knowledge and practice
    - Beliefs about student learning,
    - View of subject knowledge,
    - Use of assessment,
    - Role of teacher, student and schools
- Origin of policy pressure
- Mechanism of spread
    - Implementation plans, actual implementation or distribution of the message
- Carriers of policy pressure
- Content of policy message

BIBLIOGRAPHY

Abrams, L, M., Pedulla, J. J., & Maddaus, G. E. (2003). Views from the classroom: Teachers' opinions of statewide testing programs. *Theory into Practice, 42 (1),* 18 – 29.

Anagnostopoulos, D. (2005). Testing, Tests, and Classroom Texts. *Journal of Curriculum Studies, 37 (1), 35 – 63.*

Anagnostopoulos, D & Rutledge, S. A. (2007). Making sense of school sanctioning policies in urban high schools, *Teachers College Record*, 109 (5), 1261-1302.

Benning, V. (1999, July 5). Teachers wary of new exams: Poll finds support for D.C. program series. *The Washington Post*, p, A 01.

Blum, J. (2000, June 23). 83 schools praised for test results: Different strategies used to meet goals. The Washington Post, p. B 09.

Booher-Jennings, J. (2005). Below the bubble: "Educational triage" and the Texas accountability system. *American Educational Research Journal,* 42, 231 - 268.

Carnoy, M., Elmore, R. & Siskin, L. S. (2003). Introduction. In Carnoy, M., Elmore, R. & Siskin, L. S. (Eds.), *The new accountability: High schools and high-stakes testing* (pp.1-12). New York: Routledge Falmer.

Center for Performance Assessment. (2004). *Making standards work.* Retrieved from http://dcps.etlzone.com/

Cimbricz,S. (2002). State-mandated testing and teachers' beliefs and practice. *Educational Policy Analysis Archives*, 10 (2), Retrieved from http://eppa.asu.edu/epaa/v10n2.html.

Coburn, C. E. (2001). Making sense of reading: Logics of reading in the institutional environment and the classroom. (Doctoral dissertation.)

Coburn, C. E. (2004). Beyond decoupling: Rethinking the relationship between the institutional environment and the classroom. *Sociology of Education,* 77, 211 - 244.

Coburn, C. E . (2005) The role of nonsystem actors in the relationship between policy and Practice: The case of reading initiative in California. *Educational Evaluation and Policy Analysis,* 27 (1), 23-52.

Coburn, C.E., Pearson, P.D., & Woulfin, S. (in press). Reading policy in the era of accountability. In Kamil, M., Pearson, P.D., Moje, E & Afflerbach, PP. (Eds.), *Handbook of reading research Volume IV*. Mahwah, NJ: Lawrence Erlbaum Associates.

Cohen, D. K. (1982). Policy and organization: The impact of state and federal educational policy on school governance. *Harvard Educational Review*, 52 (4), 474-499.

Cohen, D .K. (1990). A revolution in one classroom: The case of Mrs. Oublier. *Educational Evaluation and Policy Analysis,* 12, 327 - 345.

Department of Education. (2002). No Child Left Behind: A Desktop Reference. Retrieved from http://www.ed.gov/admins/lead/account/nclbreference/index.html

Diamond, J. B (2007). Where the rubber meets the road: Rethinking the connection between high-stakes testing policy and classroom instruction. *Sociology of Education*, 80(4). 285-313.

Diamond, J. B. & Spillane, J. P. (2004). High-stakes accountability in urban elementary schools: Challenging or reproducing inequality? *Teachers College Record,* 106(6), 1145-1176.

DiMaggio, P. J. & Powell, W. W. (1983). The iron cage revised: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48(2), 147-160.

District of Columbia Public Schools. (1991). Comprehensive five year plan, SY 1991-1995. Washington, DC: Author.

District of Columbia Public Schools. (1992). The strategic plan: A five-year approach. Washington, DC: Author.

District of Columbia Public Schools. (1992). Building learning communities: A framework for preparing students for the 21st century: The District of Columbia Public Schools response to America 2000. Washington, DC: Author.

District of Columbia Public Schools. (1995). BESST. bringing educational services to students: D.C. Public Schools educational reform agenda. Washington, DC: Author.

District of Columbia Financial Responsibility and Management Assistance Authority. (1996). Children in crisis: The failure of public education in the District. Retrieved from http://www.washingtonpost.com/wp-srv/local/longterm/library/dc/control/part1.htm

District of Columbia Public Schools (1997) Moving Toward an Exemplary System: A Report on the First Year's Efforts to Reform D.C.'s Public Schools. Retrieved from http://www.dcpswatch.com/dcps/9711.htm

District of Columbia Public Schools. (2001). Children first initiative: A plan for transforming schools: A Snapshot. Retrieved from http://www.dcwatch.com/schools/ps010622b.htm

District of Columbia Public Schools. (2002). Standards for teaching and learning: Pacing charts: Ensuring success for all students! Mathematics. Washington, DC: Author.

District of Columbia Public Schools (2002). Standards for teaching and learning: Pacing charts: Ensuring success for all students! Science. Washington, DC: Author

District of Columbia Public Schools (2001). DCPS business plan for strategic reform. Retrieved from http://www.dcpswatch.com/dcps/010723.htm

District of Columbia Public Schools (n.d.) ASSESSMMrs. SommersT. Retrieved from http://www.k12.dc.us//dcps/curriculum/assess1.html

District of Columbia Public Schools (n.d.) Introduction: Using a pacing chart. Retrieved from http://www.k12.dc.us//dcps/curriculum/instruction/pacing/p-intro.pdf

District of Columbia Public Schools. (2005) Declaration of education: Keeping our promise to the district's children: Strategic plan for the District of Columbia Public Schools. Washington, DC: Author.

District of Columbia Public Schools. (n.d.) Expect the best for your child. Washington, DC: Author.

District of Columbia Public Schools (2006) Master education plan. Washington, DC: Author.

District of Columbia Public Schools. (2005) Mathematics pre-k through grade 12 standards. Retrieved from http://www.k12.dc.us/dcps/standards2005/Standards%20Guides/DCPS%20Math%20horizontal.pdf

District of Columbia Public Schools. (2006). Science pre-k through grade 12 standards. Retrieved from http://www.k12.dc.us/DCPS/Standards/General/DCPS-horiz-science.pdf

District of Columbia Public Schools. (2006). Standards-based lessons frequently asked questions. Retrieved from http://dcps.etlzone.com/

District of Columbia Public Schools. (2005). Classroom checklist. Retrieved from http://dcps.etlzone.com/

District of Columbia Public Schools. (2005). Making standards work affirmation form. Retrieved from http://dcps.etlzone.com/

District of Columbia Public Schools. (2006). Focused walks. Retrieved from http://dcps.etlzone.com/

Education Week (1003, May 12). District round up . Retrieved from http://www.edweek.org/ew/articles/1993/05/12/33dists.h12.html?qs=District+of+Columbia+Public+Schools,+payroll

Elmore, R. F., Abelmann, C. H. & Fuhrman, S. H. (1996). Holding schools accountable: Performance-based reform in education. In H. F. Ladd (Ed.). *The new accountability in state education reform: From process to performance* (pp. 65 – 98). Washington, DC: The Brookings Institution.

Elmore, R. F. (1999-2000). Building a new structure for school leadership. *American Educator, Winter,* 6 - 13.

Eisenhardt, K. M. (2002). Building theories from case study research. In Huberman, A. M. & Miles, M.B. (Eds.). *The qualitative researcher's companion* (pp. 5-35) Thousand Oaks, CA: Sage.

Everson, C. M. & Green, J. L. (1986). Observations as inquiry method. In Wittrock, M. C. (Ed.). *Handbook of research on teaching* (pp.162-213). London: Macmillan.

Firestone, W A, Montils, L F. & Schorr, R Y. (2004). The ambiguity of teaching to the test: Standards, assessment, and educational reform, (Eds.). Mahwah, NJ: Lawrence Erlbaum.

Firestone, W. A., Mayrowetz, D. & Fairman, J. (1998). Performance-based assessment and instructional change: The effects of testing in Maine and Maryland. *Educational Evaluation and Policy Analysis,* 20, 95 – 113.

Firestone, W. A., Fitz, J. & Broadfoot, P. (1999). Power, learning and legitimation: Assessment implementation across levels in the United States and the United Kingdom. *American Educational Research Journal*, 36, 759 – 793

Firestone, W. A. & Mayrowetz, D. (2000). Rethinking "high stakes": Lessons from the United States and England and Wales. *Teachers College Record,* 102 (4), 724–749.

FOSS Project. (n.d.) Retrieved from http://www.lhsfoss.org/fossweb/email/aboutFOSS.html

Glasnapp, D. R., Poggio, J. P., & Miller, M. D. (1991). Impact of a ―low stakes" state minimum competency testing program on policy, attitudes, and achievement. *Advances in Program Evaluation,* 1, 101-140.

Giddens, A. (1984). *The constitution of society*. Berkeley, CA: University of California Press.

Giddens, A. (1979*). Central problems in social theory: Action, structure, and contradiction in social analysis.* Berkeley, CA: University of California Press.

Goodstein, J. D. (1994). Institutional pressures and strategic responsiveness: Employer involvement in work-family issues. *Academy of Management Journal*, 37 (2), 350-382.

Grant, S. G. (2001). An uncertain lever: Exploring the influence of state-level testing in New York state on teaching social studies. *Teachers College Record*, 103, 398 – 426.

Hamilton, L. S., Stecher, B. M. Marsh, J. A., McCombs, J. S., Robyn, A., Russell, J. L., Naftel, S. & Barney, H. (2007). Standards-based accountability under No Child Left Behind: Experiences of teachers and administrators in three states. Santa Monica, CA: RAND Corporation.

Hamilton, L. (2004) Assessment as a policy tool. *Review of Research in Education,* 27, 25-68.

Hamilton, L, Berentd, M. & Stecher, B. (2005) Teachers' responses to standards-based Accountability. RAND Education working paper series. WR-259-EDU. Santa Monica, CA: RAND Corporation.

Hargrove, T. Y., Jones, M. G., Jones, B. D., Hardin, B., Chapman, L., & Davis, M. (2000). Unintended consequences of high stakes testing in North Carolina: Teacher perceptions. *EPS Spectrum*, 18 (4), 21 – 25.

Haynes, D. V. (2006, March 6). Janey plans to retool schools: New curricula, books, training sought by fall. *The Washington Post,* p. C 01.

Haynes, D. V. (2005, August 25). New standards will sharpen teachers' focus; City's more rigorous curriculum is designed to boost achievement. *The Washington Post*, p. T 10.

Haynes, D. V. & Labbe, T. (2006, August 24). Some revisions delayed by lack of new books. The Washington Post. p. B 8.

Hening, J. R., Hula, R. C., Orr, M. & Pedescleau, D. S. (1999). The color of school reform: Race, politics, and the challenge of urban education. Princeton, NJ:Princeton University Press.

Heubert, J.P. & Hauser, R. M. (1999). High stakes testing : Testing for tracking, promotion and Graduation. (Eds.) Washington D.C.: National Academy Press.

Hirsch, P.M. & Lounsbury, M. (1997). Ending the family quarrel: Toward a reconciliation of ―old" and ―new" institutionalisms. *American Behavioral Scientist*, 40, 406-418.

Jones, G., Jones, B., Hardin, B., Chapman, L., Yarbrogh, T. & Davis, M. (1999). The impact of high stakes testing on teachers and students in North Carolina. *Phi Delta Kappan*, 81 (3). 199 – 203.

Kauffman, D., Johnson, S. M., Kardos, S. M., Liu, E. & Perske, H.G. (2002). ―Lost at sea": New teachers' Experiences with curriculum and assessment. *Teachers College Record,* 104 (2). 273-300.

Ketter, J. & Pool, J. (2001). Exploring the impact of a high stakes direct writing assessment in two high school classrooms. *Research in the Teaching of English*, 35, 344 – 393.

Koretz, D., Barron, S., Mitchell, K. & Stecher, B. (1996). The perceived effects of the Kentucky Instructional Results Information System (KIRIS). Santa Monica, CA: RAND Corporation.

Lipman, P. (2004). High stakes education. New York. NY: RoutledgeFalmer

Madaus, G F (1988). The influence of testing on the curriculum. In Tanner, L. N. (Ed.). *Eighty seven yearbook of the National Society for the Study of Education critical issues in curriculum.* Part I 83-121.

Mathews, J. (1998, November 8) Take out your no.2 pencils: Standardized tests like the Stanford 9 are the rage in public education. But do they really measure academic achievement, or simply reflect collective anxiety about our schools*? The Washington Post*. p. W 11.

Mathews, J. & Strauss, V, (1998, November 23). Giving schools a sense of direction: Standardized tests reforming instruction in district schools. The Washington Post. p. B 01.

Mathison, S. & Freeman, M. (2003). Constraining elementary teachers' work: Dilemmas and paradoxes created by state mandated testing. *Education Policy Analysis Archives, 11,* Retrieved from http://epaa.asu.edu/epaa/v11n34

Maxwell, J. A. (2002) Understanding and validity in qualitative research. In Huberman, A. M. & Miles, M.B. (Eds.). *The qualitative researcher's companion* (37-64). Thousand Oaks, CA: Sage.

Mazzeo, C. (2001). Frameworks of state: Assessment policy in historical perspective. *Teachers College Record,* 103, 367 - 397.

Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology,* 83, 340 - 363.

Meyer, J. W., Scott, W. R. & Deal, T. E. (1981). Institutional and technical sources of organizational structure: Explaining the structure of educational organizations. In H. D. Stein (Ed.), *Organization and the human service* (151 – 179). Philadelphia, PA: Temple University Press.

McDonnell, L. M. (2004) Politics, persuasion, and educational testing. Cambridge, MA: Harvard University Press.

McDonnell M & Choisser, C (1997). Testing and teaching: Local implementation of new state assessment. Los Angeles, CA: National Center for Research on Evaluation Standards, and Student Testing (CRESST) Center for the Study of Evaluation (CSE).

Merriam, S. B. (2002). Assessing and evaluating qualitative research. In Merriam, S. B. & Associates. (Eds.). Qualitative research in practice (18-33). San Francisco, CA: Jossey-Bass.

National Council of Teachers of Mathematics (2000). Principles and Standards for School Mathematics. Reston, VA: Author.

Oliver, C. (1991). Strategic responses to institutional processes. *The Academy of Management Review,* 16, 145 – 179. Parent United for DCPS (1997). How are we doing? Look at the CTBS scores. Retrieved from http://www.dcpswatch.com/parents/pu9708.htm#How

Rowan, B. (1982). Organizational structure and the institutional environment: The case of public schools. *Administrative Science Quarterly*, 27, 259 - 279.

Schofield, J W. (2002). Increasing the generalizability of qualitative research. In Huberman, A. M. & Miles, M. B. (Eds.). *The qualitative researcher's companion* (171-203). Thousand Oaks,CA: Sage

Schneiberg, M. & Clements, E.S. (2006) The typical tools for the job: Research strategies in institutional analysis. *Sociological Theory,* 24 (3), 195-227

Scott, W. R. (1995, 2001). *Institutions and organizations*. Thousand Oaks, CA: Sage.

Scott, W.R. (2003). Institutional carriers: Reviewing modes of transporting ideas over time and space and considering their consequences. *Industrial and Corporate Change*, 12 (4), 879-894.

Scott, W.R., Ruff, M., Mendel, P.J. & Carol, C.A. (2000). *Institutional change and healthcare organizations: From professional dominance to managed case*. Chicago, IL: The University of Chicago Press.

Silverman, D. (1993). Interpreting qualitative data: Methods for analyzing talk, text and interaction. Thousand Oaks, CA: Sage.

Sloan, K. (2000, April). Teacher agency and the TAAS: Maintaining the ability to "act otherwise." Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.

Smith, M. L. (1991). Put to the test: The effects of external testing on teachers. *Educational Researcher,* 20, 8-11.

Smith, F. L. (1992) (1992, July 12). Give D.C. kids a choice; How the school chief wants to change the system. *The Washington Post*. p. C 01.

Spillane, J. (2000). Constructing an ambitious pedagogy in fifth Grade: The mathematics and literacy divide. Elementary School Journal, 100 (4),

Spillane, J.P. & Jennings, N.E., (1997). Aligned Instructional policy and ambitious pedagogy: Exploring instructional reform from the classroom perspective. *Teachers College Record,* 98 (3), 449-481

Spouvitz, J. (2009). Can high stakes testing leverage educational improvement? : Prospects from the last decade of testing and accountability reform. *Journal of Educational Change*. 10. 211-227.

Starbuck, W. H. & Milliken, F. J. (1988). Executives' perceptual filters: What they notice and how they make sense. Hambrick, D. C. (Ed). *Executive Effect: Concepts and Methods for Studying Top Managers* (35-65). Greenwich, CT: JAI Press.

Spillane, J. P. Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research*. 72(3) 387-431.

Stecher, B. M & Barron, S. I. (1999). Quadrennial milepost accountability testing in Kentucky. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESS) Center for the Study of Evaluation (CSE).

Stecher, B.M., Barron, S. I, Kaganoff. T. & Goodwin, J. (1998). The effect of standards-based assessment on classroom practices: Results of the 1996-97. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESS).

Stecher, B. M., & Mitchell, K.J. (1995). Portfolio driven reform: Vermont teachers' understanding of mathematical problem solving. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Strauss, Valerie, (May 30, 1999) School-based budget teams get mixed reviews in district. *The Washington Post.* p. C 06.

The Council of Great City Schools (2004). Restoring excellence to the District of Columbia Public Schools. Washington, DC: Author.

The D.C. Committee on Public Education (1989). Our children, our future: Revitalizing the District of Columbia Public Schools. Washington, DC. Author.

Watanabe, M (2007). Displaced teacher and state priorities in a high stakes accountability context. *Educational Policy Analysis,* 21(2). 311-368.

Weick, K. E. (1995). *Sensemaking in organizations.* Thousand Oaks, CA: Sage

Welgoren, D. (1998, April 6). Testy about D.C. exams: Some decry push to prepare students. *The Washington Post*. p. B 01.

West, P. (1994, March 23). Teacher workshops seek to make science benchmark useful tools. *Education Week*. Retrieved from http://www.edweek.org/login.html?source=http://www.edweek.org/ew/articles/1994/03/23/26aaas.h13.html&destination=http://www.edweek.org/ew/articles/1994/03/23/26aaas.h13.html&levelId=2100

West, P. (1994, September 14). D.C. gets $13.5 million NSF grant to reform math science teaching. Education Week. Retrieved from http://www.edweek.org/login.html?source=http://www.edweek.org/ew/articles/1994/09/14/02nsf.h14.html&destination=http://www.edweek.org/ew/articles/1994/09/14/02nsf.h14.html&levelId=2100

West, M. R. & Peterson, P. E. (2003). The politics and practice of accountability. In P. E. Peterson & M. R. West (Eds.), *No Child Left Behind? The politics and practice of school accountability* 1 – 20. Washington, D C: Brookings Institution Press.

Wilson, B. L. & Corbett, H. D. (1991). Two state minimum competency testing programs and their effects on curriculum and instruction. *Advances in Program Evaluation*, 1B. 7-40.

Wong, K. K., Anagnostopoulos, D., Ruthledge, S. & Edwards, C. (2003). The challenge of improving instruction in urban high schools: Case studies of the implementation of the Chicago academic standards. *Peabody Journal of Education*, 78, 39 – 87.

Zancanella, D. (1992). The influence of state-mandated testing on teachers of literature. *Educational Evaluation and Policy Analysis*, 14, 283 – 295.